PROBLEM SOLVING EXPERTISE

AND

KNOWLEDGE STRUCTURES

IN PHYSICS

A thesis submitted in partial fulfilment of the requirements
for the degree of Master of Science in Psychology
in the
University of Canterbury
by

Paul J. Tottle

1988

Acknowledgements

I am most grateful to Professor McCallion for his enthusiastic support and expert advice, and to Paul Russell, my supervisor, for his guidance with the later stages of this thesis.

Thanks to the Physics and Mechanical Engineering staff for their assistance and to the students for sparing the time and energy.

Special thanks to my father for his support at home.

CONTENTS

ABSTRACT

Recent problem solving research has focused on the distinction between experts and novices. Studies have shown that experts categorize problems according to underlying principles of the domain, while novices (possessing little appropriate knowledge) are more concerned with the surface features of the problems. This thesis confirms these findings for several levels of expertise. A further method of categorizing problems is found for naive subjects (with no apropriate knowledge). Novices also demonstrate that when required they can produce *principle* categories that are significantly different from their usual surface feature groups.

The type of knowledge that a student uses to problem solve is investigated in an endeavor to explain the low problem solving ability of first and second year physics students. The results indicate that students possess the relevent declarative knowledge, but lack the procedural form. Likewise, students do not have the ancillary knowledge required for effective problem solving. The procedures that students do have appear to be inflexible, being only applicable in a narrow range of problems. Future research can examine the naive - novice distinction in other problem solving domains (e.g. mathematics) and investigate the transfer of skills in the novice, from learning to sort by *principles* to problem solving.

CHAPTER ONE

Introduction

This thesis is concerned with the distinction between experts and novices in the specific knowledge domain of physics. The initial research sought to examine the effects of expertise on incubation in problem solving. However, methodological problems with the first experiment lead to a change of focus. Two paths were chosen for continued research: expertise and incorrect hypothesis formation. The later was unsuccessful and is taken no further than Chapter 2.The literature for these two areas will now be reviewed.

Experts and Novices

The most influential relevant work in the expertise area for the author of this thesis is a recent series of studies in the physics domain by Chi, Feltovich and Glaser (1981)[1]. Eight advanced Ph.D students (experts) and eight novices (having completed first year university lectures in mechanics) were asked to sort 24 written mechanics problems into groups according to similarities in solution. Without allowing subjects to use pencil or paper to solve the problems they found that the novices and experts produced markedly different labels for their sorted groups.

By examining the problems grouped together by more than one novice they found the problems shared certain surface features. Surface features are names for the objects in the problem (e.g. an inclined plane) and the actual words used in the problem (e.g. friction). This noticeable feature of novice groupings was not evident in the groups created by experts. The similarity of problems for expert groups could only be identified by other physists and involved the fundamental laws and principles underlying physics. For example problems using Newton's Second Law (F=ma) were grouped together.

The labels given to the groups involved surface features for novices (e.g. pulleys, springs) and principles and laws for experts (e.g. *conservation of energy*).

To test their findings Chi et al. constructed 20 new problems and tabulated them according to their *principles*[2] and surface features (study two). As expected the categories created by the experts when sorting, closely resembled the pre-determined *principle* groups. Differences did exist in the number of groups used as some experts combined pre-determined groups (e.g. *conservation of linear and angular momentum*). An advanced novice (fourth year undergraduate physic major) produced mainly *principle* group labels qualified by surface features (e.g. force problems which

[1] The research is reported again in Chi, Glaser and Rees (1982).

[2] *principle* is in italics as in the context of this thesis it refers to the fundamental principles and laws of physics.

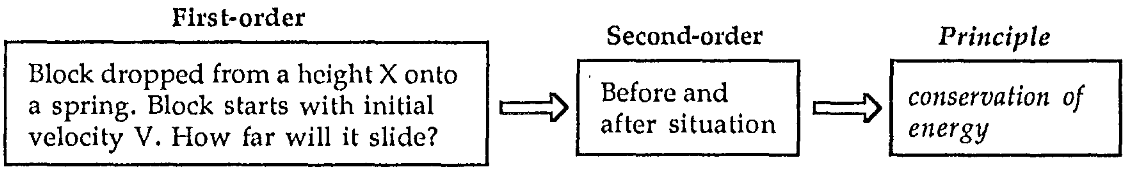


Figure 1

involve springs). The novice subject sorted by surface features. The set of problems included two specially designed problems with descriptions using energy terms, while the major principle required was, in fact, *conservation of momentum*. Both the novice and advanced novice labeled these as "conservation of energy" (the only *principle* label used by the novice). Chi et al. do not state the number of subjects used in study two, so it is likely that only the described novice, advanced novice and two experts participated.

Their third study analysed the protocols of four subjects after the subjects were asked to tell everything they could about problems that would produce 20 category labels, that had been selected from the previous studies. The published results concentrated on the descriptions given for "an inclined plane". The two experts first gave the *principles* that could be applied and the conditions for when they were appropriate (e.g. IF plane is smooth THEN use *conservation of energy*). These were followed by surface feature considerations (e.g. IF there is friction THEN put it in diagram). Only one of the two novices referred to a *principle* and this occurred at the end of the protocol. Another feature of this protocol (i.e. IF problem would involve conservation of energy and height of block,length of plane, height of plane are known THEN could solve for potential and kinetic energies) is that the conditions under which energy is conserved are left unsaid.

A further difference between the novice and expert protocols was that experts included solution methods (e.g. "use F=ma") but novices did not. This indicates that the inclined plane knowledge structures of experts schema includes explicit procedures for reaching a solution, while those of novices do not. Chi et al. point out that the experts describe a principle on the action side of a production (i.e. IF . . . THEN principle), whereas novices refer to it on the conditional side (i.e. IF principle THEN . . .). However a generality of this nature is questionable with only four subjects involved.

In study four, four subjects were asked to think out aloud about the "basic approach" that they would use as they read the problems. After each problem they explicitly stated their "basic approach" and the features in the problem description that led to it. The two experts produced almost identical "basic approaches" consisting of fundamental principles. However the novices' statements were so general (e.g. find which objects are related to each other), little could be said of them.

The feature that cued the "basic approaches" produced more interesting results. As might be expected novices referred to objects and terms explicit in the problem (e.g. spring, pulley, gravity), while experts did not. Experts identified features that consisted of conditions and states from which a "basic approach" can be determined (e.g. "no external forces", "force too complicated"). These second-order features appear to constitute intermediate states between surface features and a "basic approach" (i.e. use of *principles*). By mapping what was read and the second-order features given by subjects, the authors were able to describe the surface features (first-order) used by the experts. Generally these were large chunks of the surface features used by novices (Figure 1).

Chi, Feltovich and Glaser (1981) see "the categories of problems as representing internal schemata, with the category names as accessing labels for the appropriate schemata". (p150)

The research undertaken for this thesis tests several assumptions and possible criticisms that underlie Chi et al.'s studies. Three alternative explanations for their results are possible:

1. The expert - novice differences are due to different levels of aptitude. This assumption is tested by Schoenfeld and Herrmann ,1982 (to be discussed later).
2. Novices do not understand the problems that they are asked to sort. Therefore the expert - novice distinction is due to experts possessing more knowledge (quantity), rather than a better knowledge structure (quality). This explanation will be tested in experiment 1a of Chapter 2 (within this thesis).
3. Finally it is not clear what aspect of the problem solving process the sort task is measuring. Chi et al. assume the task is probing the initial representation of the problem. However the results may be a peculiarity of the sort process. The surface - deep distinction could do with further confirmation via an alternative methodology. Experiment 1a presents an attempt to do just that.

The well known work of Chase and Simon (1973) and the earlier work of de Groot (1966) contrasts with Chi et al.'s studies emphasizing another attribute of an expert: automaticity. By having experts and novices reproduce the positions of chess pieces on a chess board Chase and Simon were able to measure the meaningful groups ("chunks") that masters formed, and which novices did not. Newell and Simon (1972) suggested that expert chess players recognize a meaningful configuration of chess pieces and have an associated best response for each configuration. They therefore proposed that a major distinction between experts and novices was that experts had available thousands of patterns linked automatically to expert responses. Larkin, McDermott, Simon and Simon (1980) refer to this as "physical intuition".

> "A person with good physical intuition can often solve difficult problems rapidly and without much conscious deliberation about a plan of attack". (p1335)

It is ironic that at this theorized stage of development the experts' response does not require "thought". These approaches to the study of expertise have been given little space here due to their vague relationship to the concept of expert and novice knowledge structures. The meaningful *principle* categories that Chi et al.'s subjects produced in response to routine mechanics problems imply that there is more to expertise than Newell and Simon supposed. The automaticity of expert responses can be thought of as an implementation factor of "mental algorithms" (using Anderson's 1987 terminology), distinctly different from the varying knowledge structures and "mental algorithms".

Chi et al. are not alone in their research into domain-specific knowledge structure changes with expertise level. Studies have been conducted in mathematics, computer programming, video games, chemistry, baseball and dinosaur knowledge. Where these studies add to the research already presented or demonstrate the external validity of the results, they will be discussed in more detail.

Schoenfeld and Herrmann (1982) discuss the deep - surface differences exhibited by experts and novices in terms of problem perception. They designed a study to test changes in problem perception that eliminated the alternative aptitude explanation that could explain novice - expert differences (previously mentioned). That is, experts are maturer and have been through a selection process that is most likely dependent on aptitude. However, novices could have varying aptitudes as low aptitude students have not yet dropped out due to examination failure. To test this, Schoenfeld and Herrmann had the experimental and control groups perform a mathematics card sort (similar to Chi et al.'s sort) followed by a mathematics test. For the next month the experimental group participated in a class teaching techniques in mathematics problem solving (e.g. general heuristics). The control group were enrolled in a computer science course, "Structured Programming". At the end of the month the card sort was repeated for all subjects followed by another mathematics test.

By performing t-tests and correlating the before and after sort matrices the researchers showed a significant shift from surface to deep structure by the experimental group. The group also showed a marked improvement in the mathematics test. So by using the same subjects at different levels of expertise over time, Schoenfeld and Herrmann have removed the confounding effects of age and aptitude.

Means and Voss (1985) had interesting results when investigating age differences with expert and novice knowledge structures. They used the knowledge domain of the movies "Star Wars" and "The Empire Strikes Back". Subjects were asked a series of detailed questions to extract information from various levels of a pre-determined hierarchical structure. For example, at a high level there were the goals of the Rebel Alliance while the lowest level consisted of the actions undertaken by the characters at different times. Using second to ninth grade school children they found that at a certain level of expertise (measured by the number of viewings of the two movies) there are age related knowledge structure differences.

As may be expected older experts were able to answer the questions in more detail showing a quantitative difference with age. Qualitative differences were found to exist in the high level structure of the hierarchy. Younger experts talked about good fighting bad, while older experts also referred to political and ethical features. Means and Voss briefly interpreted their findings in terms of schema theory. Older experts have an "international conflict" schema and younger experts a "good guy - bad guy" schema. These schemas play an active role in the interpretation and coding of the "Star Wars" dialogue. However this is almost a chicken and egg situation if it is supposed to explain the "Star Wars" knowledge structure development.

As the researchers point out, the "qualitative" differences at a high level could have created the quantitative age differences. Older experts have a better discriminating hierarchy (with more components), allowing for more detailed storage and retrieval.

The interesting finding of this study is that the knowledge of the younger experts appears to be a subset of the older experts, rather than involving a different knowledge configuration (e.g. surface - deep structure). However due to the authors establishing a goal oriented knowledge structure and using cause and effect oriented questions (e.g. "Why ...") it is not clear as to whether a surface - deep structural difference exists but goes undetected.

By including characteristics of the problem situations, declarative knowledge, and procedural knowledge in a sort task, de Jong and Ferguson-Hessler (1986) set out to show that good novice problem solvers have their knowledge organized according to problem schema (and poor novice problem solvers do not). The novices were selected from a first year university course on electricity and magnetism. They were divided equally into two groups according to the mean score in the course examination. The researchers found a strong correlation between the examination results and the sort matrix (i.e. 65 cards produced a 65 x 65 matrix of pairs of cards that were grouped together).

Through examination of the elements of each group formed, they found that the good novices sorted by the "problem-type" groupings established by the researchers. The poor novices however formed groups characterized by words on the cards (e.g. "fields", "calculate", or "energy and work", and "w").

The earlier work of Silver (1979) and Chartoff (1977, cited in Silver, 1979) used two different approaches (sorting and multi-dimensional scaling) to study the relatedness of mathematics problems and came to similar conclusions. They found that subjects judged problem similarity on two major dimensions "mathematical structure" and "contextual details". Silver showed the "mathematical structure" variable to be correlated with various measures of mathematical ability.

Hypothesis Theory and Problem Solving

The important aspects of problem solving, the interpretation and the creation of a particular initial representation emphasized by previous researchers, can be studied through the examination of the hypothesis process (formation and evaluation). Three studies will be explained, that indicate the potential importance of early hypothesis formation to problem solving.

Bruner and Potter (1964) presented subjects with a series of complex pictures that were either very blurred, moderately blurred or slightly blurred. The pictures were slowly focused to a point at which a high percentage of subjects could normally identify the picture correctly. Subjects attempted to identify the pictures as they were focused. The researchers found that at the same focal length the subjects in the slightly blurred group did significantly better than those in the very blurred group. By examining the subjects protocols they found that hypotheses were made early on despite the lack of information. This hypothesis then needed to be disproved before a better hypothesis could be considered.

> "The ambiguity of the stimulus is such that no obvious contradiction appears for a time, and the initial interpretation is maintained, even when the subject is doubtful of its correctness". (p425)

Another study (Levine, 1971) showed a non-learning effect of early, incorrect hypothesis formation. Subjects were required to identify the correct card of two presented (Figure 2).

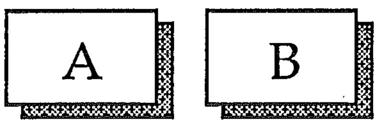


Figure 2

Whenever the subjects said "A" they were told "right" and whenever subjects said "B" they were told "wrong". As one would expect this simple reinforcement contingency task was quickly solved. However Levine preceded this task with a more complex sequence "consisting of a cycle of 10 trials with five shifts during a cycle" (e.g. left-left-right-left-left-left-right-left-right-right). Only 10 of the 52 subjects managed to solve the simple task after over 100 trials. Feedback from the subjects that did solve the problem indicates they were actively looking for a "trick" to the experiment. In terms of Hypothesis theory subjects were sampling from a set of hypotheses which do not contain the correct hypothesis.

Hypothesis theory has three basic assumptions:

1. A subject selects an hypothesis from some domain and responds on the basis of that hypothesis.
2. The domain from which the hypotheses are sampled is delimited by the stimulus information.
3. If when the hypothesis is tested it is found to be wrong the domain is re-sampled with the set of hypotheses reduced in size. (The extent of the reduction has been a point of great debate).

Due to the use of Einstellung (set effect) and the sequence effect in the methodology of experiment 1a (Chapter 2), Sweller and Gee's (1978) extension of Hypothesis theory to cover these, will be considered. Briefly set effect is produced when subjects have an inability to solve a simple problem after solving a series of different problems. The sequence effect occurs when subjects can more rapidly solve an easy-to-complex series of problems than a complex-to-easy sequence. To readily account for these two effects the authors proposed two further Hypothesis theory criterion:

1. "When subjects solve a series of problems that they perceive as being related, they begin each problem by testing hypotheses as closely related as practicable to their previously correct hypothesis".
2. "There is a positive relationship between the complexity of a hypothesis and the number of hypotheses or rules that are related to it".

Experiment 1a of this thesis uses both the existence of Einstellung and the sequence effect in an attempt to test the generality of Chi et al.'s findings. In terms of Sweller and Gee's relatedness of hypotheses, presumably for experts the relatedness refers more to deep structure relations and surface relations for novices.

The Present Research

The two Hypothesis theory experiments (Bruner & Potter, and Levine) were presented to illustrate the possible harmful effect of subjects forming hypotheses during the reading of a problem, rather than waiting until they had all the relevant information. A possible expert - novice distinction is that experts do not form premature hypotheses, while novices do. Experiment 1a (Chapter 2) attempted to test this by reversing the question order so that the actual question is stated before the background information is given. That is, to encourage a more working backward / means-ends-analysis approach.

To test the generality of Chi et al.'s findings a more complex method was used. It was based on the following assumptions:

1. Experts have their knowledge organized (i.e. categorized at the top of the hierarchy) by underlying principles. Novices have their knowledge organized by surface features.
2. A high proportion of top down processing of the input results in experts detecting, identifying and categorizing fundamental principles first, while novices detect, identify and categorize surface features first. Therefore, experts will detect a change in principle usage from one problem to the next faster than a novice.

It was expected that presenting a series of surface and *principle* related problems followed by a different problem would induce a set effect. However if this latter problem was only different in *principle* (i.e. it still looked the same) then the novice would be unaware that a change in method was required, and fall for Einstellung. The expert who codes by *principles* would see a difference in the problem and sample a different domain.

CHAPTER TWO

Research

Experiment 1a

The objective of the first experiment was to determine the influence of the surface/deep sort structure difference reported by Chi, Feltovich and Glaser (1981). The difference between experts and novices was studied in a repeated measures design involving physics mechanics problems. One group of subjects solved three Newtonian problems followed by one requiring a *conservation of energy* solution method. All four problems contained the same surface features. The second group of subjects solved only the *conservation of energy* solution problem. It was expected that the difference in solution time for experts between the two groups would be smaller than for novices, as experts can quickly overcome the set effect induced by the Newtonian problems due to a *principle* organized knowledge structure.

To test the disadvantage of early hypothesis formation a further group solved the problems in which the problem question was written before the problem description. Therefore it was expected that novices would solve the problems quicker with the question first. While the effect would be far less noticeable with the experts.

METHOD

Subjects

Twenty two first year (not repeating) physics students (20 males and 2 females) participated in the experiment. All subjects had completed 2-3 years of physics at high school as well as completing the mechanics section of the PHYS 101 (1st year university physics) paper. The students in the most homogeneous (determined by total Bursary points) laboratory class of PHYS 101 were individually asked to partake. Fifty students were spoken to, seven did not wish to help while thirty six of the remaining forty three kept their appointment (including experiment 1b). The experiment was timed to follow the mechanics lectures yet precede their build up to the mid-year examination. That is, to avoid a confounding practice effect over the days of the recordings (caused by the students preparing for their examination).

Materials

Four questions were selected from the textbook used by the students; Ohanian (1985). They were modified so that each question had the same surface structure. Diagrams showed two masses on a pulley system, with one mass a height x above the ground. The questions retained the same word order where possible. Problems 1 and 2 referred to one diagram while problems 3 and 4 referred to another. However the two different diagrams contained the same surface features. (Two diagrams were used to provide some of the variety in illustrations students are accustomed to). The questions are presented in full in Appendix A.

The first three problems (P1, P2, P3) asked for calculations based on Newton's second law (in its simplest form: Force = mass x acceleration). The final question (P4) asked for the energy gained by m_2 . This can be calculated quickly using the *conservation of energy* (a complete change in principle to that of Newtonian physics[1]).

A second series of problems (P1' ,P2' ,P3') were identical to P1, P2, P3 except the actual question was asked before the details were given. For example, question 2:

> What is the upward acceleration of the elevator cage if the pulleys are permitted to run freely? (Derive a formula for the acceleration and solve)
>
> Given that an empty elevator is an unknown height above the ground and consists of an elevator cage of mass m_1 (1200kg) connected by a cable, running over a pair of pulleys, to a counterweight m_2 (1300kg). Neglect the mass and friction of the cable and pulleys.

Subjects were provided with a simple calculator as well as pen and paper.

Procedure

Subjects (seen individually) were first asked a series of questions to determine their physics experience and ability (i.e. number of years of school physics, when did they last solve a physics mechanics problem, whether they are repeating the PHYS 101 course and their Bursary physics mark). Following this, the students were given a copy of the instructions.

> The following mechanics problems have been adapted from the physics text book used by PHYS 101. One question must be completed before moving onto the next question. I will be recording the time required to complete each question, so please tell me when you believe you have finished each one. Give all working out in full as you would in an examination, plus any relations / formula you are contemplating but may not use. Please write in order down the page. Are there any parts of these instructions that need explaining?
>
> For the following problems the acceleration of gravity = 9.8 m/s^2.

The researcher elaborated on the instructions emphasizing that subjects report when they believed they had finished.

The experimental Group-1 solved P1, P2, P3, P4 while the Control group solved only P4. The experimental Group-2 were asked to solve the three rearranged problems (P1', P2', P3'). Subjects were randomly distributed between the three groups. The problems were presented separately, in the order: P1 ,P2, P3, P4. Only one problem was available at a time. Subjects were timed from the presentation of each problem until they told the researcher they believed they had completed the answer. The timer was stopped and the student's solution checked. If the result was incorrect the subject was told so and asked to continue with the problem. The time was recorded, and the stopwatch restarted (the combined time was recorded). If after an interval the subject did not believe he had the ability to complete the problem or the time exceeded 10 minutes, without the subject nearing the solution, the attempt was aborted. Subjects were then given a model answer and allowed as long as they wished to examine it. The answer was left on the table available for the subjects reference in later problems. The researcher recorded any comments made regarding why difficulty had been experienced in solving a problem. Finally the researcher stressed the importance of not mentioning the contents of the experiment to anyone.

RESULTS

A serious problem became evident after testing only 4 subjects. This resulted in the experiment being redesigned (see "Further Results"). The trouble arose quite simply from the inability of subjects to solve P4. The first three subjects failed to solve the problem. They all applied energy theorem to the problem, but did so incorrectly. Two of the subjects tried to combine Newtonian and

---

[1] For this thesis Newtonian physics refers specifically to Newton's laws of motion.

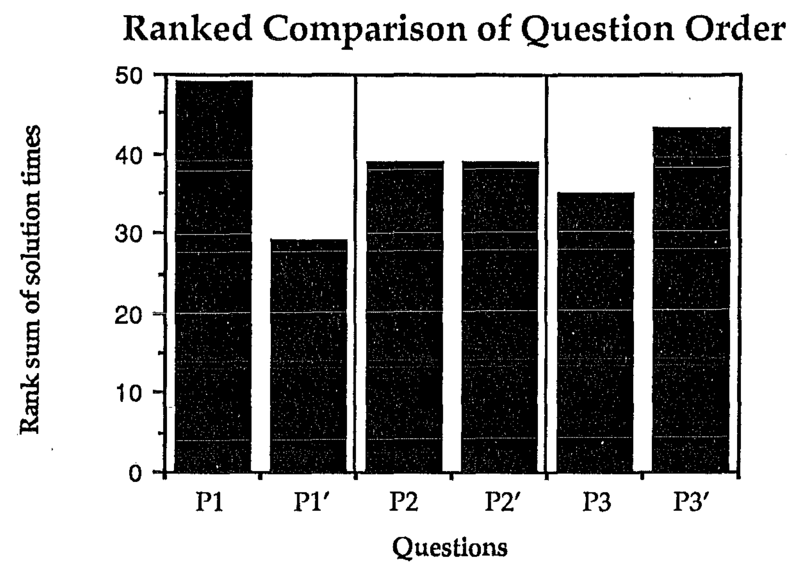


Figure 3

energy methods unsuccessfully (though this is possible). The fourth subject solved the question in 11 minutes using Newton's laws of motion (i.e. using the same principle for all four problems).

The difference in solution time between the question order groups (Group-1 & Group-2) was not in the predicted direction for P2/P2′ and P3/P3′ (Figure 3). The graph displays the sum of the ranked values used in the calculation of theWilcoxin Rank-Sum test. A Wilcoxin Rank-Sum test of P1/P1' gave an insignificant value of 28 ($p[U \leq 28]=0.07$). The non-parametric test was necessary due to the failure of some subjects to solve a problem in the set time. The minimum solution times for the problems were 2.5 (P1), 2.3 (P1'), 3.3 (P2), 3.3 (P2'), 5.5 (P3), and 9.35 (P3') minutes. The maximum solution time needed for P1' was 6.1 minutes while the other questions were not solved by at least one subject.

Discussion

The results from the question order conditions (Group-1 and Group-2) did not support the researchers prediction (based on Hypothesis Theory). That is, the results do not support the concept of early, disadvantageous hypothesis formation by the novices. Because of the surface feature similarity of the problems (P1'-P3' or P1-P3), it was expected that the effects of presenting the actual question before the detailed problem description would decrease after the first problem (i.e. subjects would remember the content of the previous question). However the P2/P2' and P3/P3′ differences are not in the direction predicted.

The failure to solve P4 produces a series of unanswered questions: What do the novices know?; Did Chi et al.'s subjects understand the problems they were sorting?

Further Results

Problem four was put through a series of exploratory changes (Figure 4, over the page) in an attempt to discover what students knew and to find a workable question.

a) How much kinetic energy does the mass m_2 gain by the time the man lands on the ground? (Derive a formula for the energy without using kinematic relations e.g. $v^2 = u^2 + 2as$).

Variations in P4

	Original	a	b	c
Subjects	4	4	4	4
Correct	1	0	0	2

Figure 4

The initial energy of the system is equal to the final energy of the system, so:

$$\frac{1}{2}(m_1 + m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1 + m_2)v_f^2 + m_2gx$$

$$v_i = 0 \qquad \text{therefore} \qquad v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1 + m_2)}$$

Gain in energy of mass m_2 when the man lands is:

$$= \frac{1}{2}m_2v_f^2 + m_2gx$$

Substitute the equation for v_f^2 into this equation . . .

Figure 5

The previous working out is reduced from the 4 steps shown above to the first 2 steps and K.E. (Kinetic Energy) = $\frac{1}{2}mv^2$ (Figure 5).

However it quickly became evident subjects were having difficulty applying the *conservation of energy*. That is, they were unable to produce the correct theorem (line 1 in Figure 5). Most subjects simply used $\frac{1}{2}mv_1^2 = m_2gx$, neglecting the kinetic energy of mass m_2, despite the explicit reference to kinetic energy.

The inclusion of "without using kinematic relations e.g. $v^2 = u^2 + 2as$" eliminates the solution path by Newtonian physics ($v^2 = u^2 + 2as$ is required). In terms of Hypothesis Theory the search domain has been either reduced or changed by this statement, resulting in reduced solution time. Nevertheless, one subject persisted with Newtonian physics for 15 minutes. None of the four subjects was able to complete this question correctly.

b) With what speed does the man hit the ground? (Derive a formula for the speed without using kinematic relations e.g. $v^2 = u^2 + 2as$).

The explicit surface reference to energy is removed with this question. The solution still remains the same (i.e. The first 2 lines of Figure 5). The change in solution approach is dramatic. The four subjects each tried to apply Newtonian physics and failed to solve the problem. One subject included an incorrect statement of the *conservation of energy* with his Newtonian approach.

c) i) What is the change in potential energy in the system ? (a gain or loss?)

ii) What is the total kinetic energy of the system as the man reaches the ground?

iii) How much kinetic energy does the mass m_2 gain by the time the man lands on the ground?

This format (common in 1st year physics questions) guides the student through the correct application of the conservation of energy principle (i & ii) but still requires students to have an overall perception of the problem (to answer iii). Students are directed to consider the energy of the whole system, rather than of the particular masses. The answer is shown in Figure 6 (over the page).

All four subjects solved part i & ii. Two of the subjects solved part iii. The average times for the three parts were; 4.05, 0.63, 5.32 (2) minutes.

i) There is a loss of P.E. $= (m_2 - m_1)gx$

$= \ldots\ J$

ii) As there is no loss of energy in the system due to friction...
the loss in P.E. must now be a gain in K.E. (*conservation of energy*)

K.E. $= (m_2 - m_1)gx$

$= \ldots\ J$

iii) This can be calculated by looking at the mass proportions ...

$$\text{K.E. of } m_2 = \frac{m_2}{m_1 + m_2}(m_2 - m_1)gx$$

or alternatively:
The initial energy is equal to the final energy, so:

$$\frac{1}{2}(m_1 + m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1 + m_2)v_f^2 + m_2gx$$

$$v_i = 0 \qquad \text{therefore} \qquad v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1 + m_2)}$$

$$\text{K.E. of } m_2 = \frac{1}{2}m_2v_f^2$$

Figure 6

Discussion

Only 3 of the 16 subjects solved the problem. It can be concluded that in general the students did not have the correct procedural knowledge necessary. However due to the extensive use of the *conservation of energy* theorem (12 of the 13 incorrect) students are able to identify the correct principle. A declarative knowledge structure exists for the principle (though its application is far from correct). Despite Chi et al.'s findings that novices sorted by surface features, the students had no trouble initially generating the principles. This leads one to wonder if the sort task is probing the initial representation of the problem. If it isn't, what is it probing? Could it in fact be measuring a feature that is unique to the sort task? If this is so, the interpretations made by Chi et al. are questionable.

It can also be argued that the principles written by students are part of a later stage of problem solving produced after the initial representation. In which case the results here do not counter the view that the sort task in measuring the initial representation. However the results are now open to more criticism.

Student's first step in their problem solving procedure was to state the principle e.g. $\frac{1}{2}m_1v_f^2 = m_2gx$. However this rigid procedure is far from appropriate. An interesting question is whether a student has the declarative knowledge to solve the problems but not the procedural knowledge. The fact that subjects solved problem four suggests this is so. The idea of a differential development of declarative and procedural knowledge is taken up later (experiment 2). A closer look at P4c and its answer shows that i) and ii) by-pass the mistake made previously, of not considering the K.E. of m_2. Part iii then explicitly asked for the K.E. of m_2. These are the points within the subjects procedure that are apparently not flexible enough to take into consideration the variations in problem (e.g. a velocity for m_2 when the velocity of m_1 equals zero).

Hypothesis Theory and Problem Order

In hindsight, there are several methodological problems with the literature and present research that could account for the failure to show a significant effect with problem statement order. Presenting the question first assumes it contains the crucial information and is sufficient on its own to elicit a correct hypothesis formation. However this is not so. The remaining information is still required to determine which is the appropriate procedure. If subjects possessed only one procedure for determining the velocity, acceleration, etc. no further information would be needed. This could in fact explain the significant result for the first problem, "What are the tensions in the cable ...?

That is, novices at their particular stage of experience may possess only one procedure for calculating tensions. The subjects of Experiment 1 would definitely have more then one procedure for calculating velocities or accelerations.

One of the research papers discussed in the introduction (Chapter 1) has a possible flaw. Bruner and Potter (1964) had subjects report aloud their interpretation of the blurred pictures. However Hislop and Brooks, 1968 (cited in Posner, 1973) found a strong effect of verbalizing an answer on hypothesis formation that also cautions the use of protocol analysis. Subjects were shown cards with cartoon animals on, that varied on several dimensions (e.g. colour, type of animal). As with the Levine (1971) study subjects had to say whether a card did or did not comply with an unknown rule (in this case, cards with two or more of the same animal). They found that a group that was required to verbalize the rule before classifying a card produced significantly fewer correct codings than a group that classified first. "The authors concluded that the subjects who had to verbalize first tended to let their behaviour be governed by their verbal hypotheses". (p75)

If this is the case, Bruner and Potter's finding could also be due to a tendency to be committed to a behavioural course because of an early verbalization. The issues relating to verbalizing a response are discussed in more detail in Ericcson and Simon (1984).

The basic problem with the problem order research undertaken here is methodological which could possibly be overcome through the use of a simpler problem domain. As Chapter 3 will illustrate, this is now only a sideline to the main research of this thesis and will be pursued no further.

Experiment 1b

To make any further re-evaluations of Chi et al.'s findings it is necessary to determine whether: New Zealand's 1st year university students (novices) sort by structural properties; and more importantly do novices understand the problems they are asked to sort. To test this latter point students were simply asked to sort by fundamental laws or principles of physics. Students were asked to sort by similarity in solution method instead of by similarity in solution (Chi et al.) to determine the importance of the task description. It was expected that students would provide more deep structure categories with the more explicit sort by similarity in solution instruction.

METHOD

Subjects

Thirteen first year (not repeating) physics students originally set to participate in Experiment 1a took part in this experiment. Subjects were selected from the same population as Experiment 1a, that is, they had completed 2-3 years of school physics and the mechanics section of PHYS 101.

Materials

Twenty one physics problems were selected to match Chi et al. (1981)'s description of their sort task (Figure 7a). Following Chi et al., the problems were not accompanied by diagrams. They were taken from the three introductory physics texts (Kane & Sternheim, 1983; Ohanian, 1985; Sears, Zemansky & Young, 1982) available to students at the reserve desk in the university library. Figure 7b shows the problem types chosen[2]. Problem 3 and problem 4b (speed) from Experiment 1a were included. The wording of any questions that originated from the main text in one of the books was changed to reduce any direct retrieval by the student of the topic involved. The questions were printed and pasted onto 8cm by 11cm cards. Each card had a number in the top left corner coinciding with the numbers in Figure 7a and 7b. The questions are printed in full in Appendix B.

[2]Unfortunately the Chi et al. authors did not reply to a letter asking for a copy of the 20 problems they designed. Therefore problems were chosen with Figure 7a as a guide-line.

Procedure

After collecting demographic information on physics experience, subjects were asked to sort the problems on the cards into groups by either their "fundamental laws or principles of physics" (5 subjects), or by "similarities in solution method" (4 subjects), or by "similarities in solution" (4 subjects). Subjects were told they would need to make their own interpretation as the researcher could not elaborate on the description. The 'sort by' instruction remained visible throughout the task. The students were also instructed that if a problem was different from all other problems it was quite acceptable to place it in a group of its own. Therefore the extreme example is 21 groups for 21 cards. Alternatively all 21 cards may be similar and placed in one group. Subjects were told that they were being timed only to give an indication of how much time was being spent on the task. They had as long as they wished to complete the sorting. It was also mentioned that the numbers on the cards bore no relation to the task at hand.

On completion of the exercise the researcher recorded the card numbers and groupings, together with the subject's explanation of what were the common factor(s) holding the cards in a group. On most occasions a description was also taken of how the 'sort by' instruction was interpreted (i.e. what characteristic was used to discriminate between the groups).

RESULTS

Questions 9 and 17 were dropped from the analysis as they were incorrectly written and created unsolvable problems.

Surface Features	Forces	Energy	Momentum (Linear or Angular)
Pulley with hanging blocks	11 14*	20- 19- 3*-	
Spring	18	7 16 9	1 17+ 6+
Inclined plane	14*	3*- 5	
Rotational	15		2 13
Single hanging block	12		
Block on block	8		
Collisions (Bullet-"block" or Block-block)			4 6+ 10+

* Problems with more than one salient feature. Listed by each feature.
- Problems that could be solved using either of two principles, energy or force.
+ Two-step problems, momentum plus energy.

Figure 7a

Surface Features	Newton's 2nd Law	Conservation of Energy	Conservation of Momentum (Linear or Angular)	Principle of Moments	Hook's Law
Pulley with hanging blocks	11 14*	19- 20- 3*- 23			
Spring		7 1 16 9	17+*		18
Inclined plane	14*	3*- 5			
Rotational			2 13	15	
Single hanging block					
Block on block	8				
Collisions (Bullet-"block" or Block-block)		12	4 22 21+ 17+* 10+		

* Problems with more than one salient feature. Listed by each feature.
- Problems that could be solved using either of two principles, energy or force.
+ Two-step problems, momentum plus energy.

Figure 7b

An example sort grouping for each of the three instructions is given below. An attempt was made to show a grouping that is typical for each instruction, however the *solution* and *solution method* groupings appear the same (in terms of *principles* and surface features). Sort by *solution* :

"spring"
"pulleys with masses interrelated somehow"
"collision"
"inertia / rotation"
" ? "
"sort of related to pulley"
"friction"

Sort by *solution method*:

"momentum"
"motion of rigid bodies"
"frictions"
"springs"
" ? "
"forces"

Sort by *principle*: .

"springs - forces F=ky"
"radians - speed (rad s^{-1})"
"collisions"
"F=ma - forces and gravity"
"energy conservation"
"friction"

(The bold print shows the group labels coded as *principles*. A "?" indicates a group of problems that the subject could not classify).

The mean solution times for the three sorting requests were 16.93 (by solution), 11.18 (by solution method) and 16.10 (by laws or principles) minutes. Aposteriori two-tailed t-test revealed that the subjects did not sort significantly faster when sorting by *solution method* .

A method of coding the results was devised to give a measure of the use of principle groupings and the correctness of these. Group labels were initially coded as *principle* or *surface*[3]. Groups with suspect labels (e.g. energy) were checked by the researcher and coded to *principle* if the majority of the cards required the related principle (e.g. *conservation of energy*) to solve. The problems on several cards could be solved using either Newtonian Physics or *conservation of energy,* however in most cases one of the methods was considered incorrect due to its complexity[4]. Problems that required both *conservation of momentum and energy* to be considered were coded as correct if either was used as the group label (According to Chi et al. experts tended to create another group for problems requiring two principles). If a subject had several groups with the same *principle* but subgroups by surface features these were counted as only one *principle* label (e.g. spring-Hooks Law + projectile-Hooks Law). The following descriptive statistic was used for each subject:

$$P = \frac{\text{number of } \underline{\textit{principle}} \text{ labels}}{\text{total number of labels}}$$

A further statistic was also calculated:

$$C = \frac{\text{number of correct cards in the } \underline{\textit{principle}} \text{ groups}}{\text{total number of cards in those groups}}$$

The proportion of *principle* labels is illustrated by P, while the proportion of correct cards within these labels is given by C. The statistic C did not provide any additional information and therefore was used no further. The results for each subject of the P statistic is presented in Appendix C. Figure 8 over the page, shows the frequency of each P value for sorting by solution and *principle.*

Due to the low variance and skewed nature of the data, the following nonparametric test was used: Wilcoxin Rank-Sum test, normal approximation, corrected for ties following formula 2.2 for variance and 2.3 for correction of ties, from Leach (1979, pp.49-73). The *solution method* group did not prove significantly different from the other two groups. However the *solution* and *principle* difference for P (Figure 8) proved significant ($p<0.05$).

If the number of cards correctly placed in

---

[3]Professor McCallion (Mechanical Engineering, University of Canterbury) helped considerably by coding the novices' group labels and explaining the possible solution methods.

[4]The solution methods described by Professor McCallion as "hellish"(!) were not considered acceptable responses from novices.

Frequency of P values

	P values 0	0.17	0.25	0.33
Principle	1	0	1	3
Solution	3	1	0	0

$S_c = -13.17$
$z = -1.72$
$p < 0.044$ (one tail)

Figure 8

priniple groups, totaled for all subjects is divided by the total number of cards sorted for all subjects, the result as a percentage represents the proportion of cards correctly coded by the students. For the *principle* sort 28.6% of the cards were correctly coded.

Discussion

The *principle* sort results were significantly different from the sort by solution results suggesting that these two types of instructions are generally treated as different questions. It can be seen that the interpretation the novices put on the initial sort by instruction played an important role in the labels produced. The vagueness of the solution sort instruction was shown in the diverseness of the group names. For example, one person sorted the cards into groups of similar complexity.

This result shows that students do have the knowledge of *principles* necessary to produce *principle* groups when asked to sort by solution. However the results show that only 28.6% of the cards were correctly coded under a principle label when asked to sort by *principle.* That is, subjects either do not know what fundemental principles or laws are or they only understand 28.6% of the questions. If students do not understand the questions, the large expert/novice discrepancy could be due more to a lack of knowledge on the part of novices than structural differences.

Discussions with Professor McCallion as he solved several of the card problems , plus the results from the novices described above, lead to the development of a simple working model (Figure 9).

Professor McCallion believes he first accesses his "belief system" about the conditions under which the problem exists. Following this a method for solving the problem arises. Novices however appear to have a rigid set of procedures in which the declarative knowledge has little or no influence.

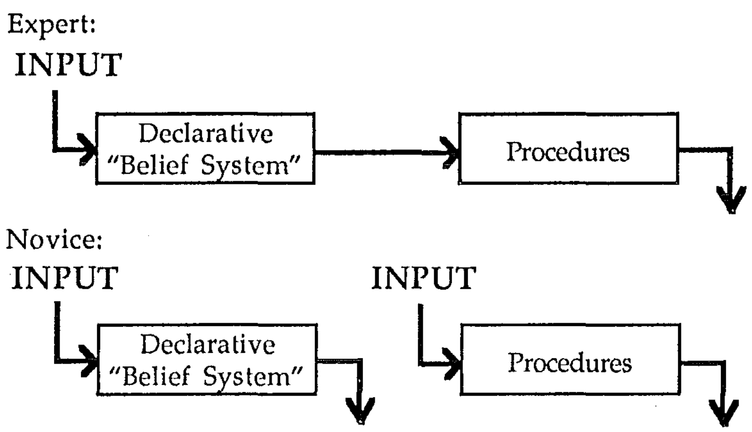


Figure 9

Experiment 2

The following experiment was designed to show if subjects have information available in declarative form in addition to that used in procedures. If as Professor McCallion proposed, experts access declarative knowledge first which then indicates the appropriate procedure to apply, novices could be expected to have more success if they are made to access declarative knowledge first. This is based on two basic assumptions: Experts have a more efficient knowledge structure for solving simple physics problems and novices will benefit (without practice) by using this strategy. Unfortunately the proximity of a class test and the inability of the few subjects who did attend to solve the problems resulted in the experiment being a pilot study.

METHOD

Subjects

Six male 1st year professional engineering students, having completed PHYS 101 in a previous year, arrived at the appointed time. One further student had attained direct entry in to the engineering course (by-passing the PHYS 101 requirement), and was therefore not included. All students attending one of the two engineering design laboratories were asked to partake. Appointments were made with twelve of the twenty students there.

Materials

Three questions were designed to give an indication of a subject's declarative knowledge (Figure 10a). These were later changed to the form shown in Figure 10b. Further to these questions, problem 4b from experiment 1a was used with an explicit reference to the *conservation of energy*:

> With what speed does the man hit the ground? (Derive a formula for the speed by considering the **conservation of energy**).

The sort cards described in experiment 1b were also used (Appendix B).

Instructions
Please answer the following three questions as you would in an examination; expressing your ideas as fully as possible, to show an examiner you know what the area involves. I will be timing you, but this is just to give me an indication of how much time you spend on the task. It may take about 15 minutes.
There is one more problem to answer after these three.

Questions
1.Explain in one paragraph the principle of *conservation of energy* and its implications.

2.In the context of mechanics, what is tension ?

3. Explain how the following two groups differ in terms of their role in physics.

torque	conservation of energy
inertia	principle of moments
rotation	conservation of momentum
velocity	Newton's laws

Figure 10a

Questions

1.Explain in one paragraph the principle of *conservation of energy* and its implications in the context of the problem.

2.Explain in the context of the problem, what is tension

3. Explain how the following two groups differ in terms of their role in physics.

torque	conservation of energy
inertia	principle of moments
rotation	conservation of momentum
velocity	Newton's laws

Figure 10b

Procedure

The experiment consisted of two treatment groups. Group-1 answered questions 1-3 followed by the problem. Group-2 were asked to solve the problem then they performed the sort task. Before solving the problem subjects were instructed as with experiment 1a, to treat the problem as they would in an examination, answering in full. They were told they would be timed and therefore it was important they indicated when they believed they had finished. Students were asked to try again if they gave the wrong answer and the stop-watch was restarted (the combined time was recorded).

RESULTS

Students in both groups, despite having completed a course in first year physics, failed to solve problem 4a. Problem 2 and 3 of experiment 1a were also given without any successes. The same mistakes were made in each treatment as those made by PHYS 101 students in experiment 1a (e.g. $\frac{1}{2}m_1v_f^2$ equated to m_2gx in problem 4 and m_2 left out of the final equation in problem 2).

The answers given to the questions (Figure 10a) did not show any greater understanding than shown in the problem solving methods of experiment 1a. Question 1 was answered by stating the basic principle: energy can be transfered but not created or destroyed. When refering to the problem (Figure 10b) no more was added than was stated in the problem. For question 3 students basically answered correctly, expressing that the members of the second group are theoretical and made up of the first group: "inertia, rotation, velocity and torque", which are physical terms, measures or quantities. One student incorrectly wrote that the first group refered to angular motion and the second to linear motion.

The questions were made more specific (Figure 10b) but most subjects did not mention the missing information that the whole system (comprising both masses) must be considered throughout the energy theorem.

Due to the time restrictions only two subjects performed the "sort by similarity in solution" task. One of the two students used a ***principle*** group.

Discussion

Unfortunately the major hypothesis, that novices would solve problems quicker after assessing declarative knowledge, was not tested.

The students demonstrated they knew the basic facts about the conservation of energy, however they were unable to provide any specific facts within the context of a particular problem. This simple result suggests there should be more educational emphasis on the procedural knowledge (problem solving) compared with the theory.

CHAPTER THREE

General Discussion

Low Problem Solving Ability

The surprising result of experiment 1a was the inability of first year physics students to solve simple mechanics problems after a course and tutorials on the subject. First year university physics does not appear to give students an understanding (a knowledge structure) for future problem solving. It appears to familiarize students with the topics and laws which they may have rote learned for examination purposes (though there is no clear evidence of this). In most cases the students were able to identify the principle involved but were unable to apply an error free procedure. The procedures they applied were often inflexible (not changing to the specific requirements of the problem) and inappropriate. There are several avenues that the research of experiment 1a can progress along:

1. To continue with the methodology of experiment 1a the study of subjects with a higher level of expertise (experience) is required. Experiment 3a follows this line of enquiry.
2. Researching the differences between experts and novices will indicate the knowledge that experts have and novices do not. As the research reviewed in the introduction made apparent, these differences are both qualitative and quantitative.
3. Determining what novices and experts know about the appropriateness and correctness of specific problem solving procedures. A model is proposed in a later section to account for flexibility and appropriateness.

The Sort Task

Following a more qualitative line, experiment 1b successfully replicated the work of Chi et al. and others, producing a surface sort structure for novices. However the results both expand and question the conclusions of past research. It was shown that a change in the instructions (solution *to* fundemental principles or laws of physics) produced a change in the novices response. One could propose that the knowledge structure difference between experts and novices is due solely to the level of interpretation of the task instructions. That is, novices are capable of interpreting the sort instructions as experts do, but chose not to. This serves as an explanation for the results. However the fact that experts do sort by *principles* despite their not being explicitly asked to, demonstrates an interesting distinction from novices, that requires further investigation.

The finding from experiment 1b that novices can sort by fundemental principles or laws of physics, adds an interesting dimension to Chi et al.'s results. Even though the students demonstrated they were unable to solve the problems of

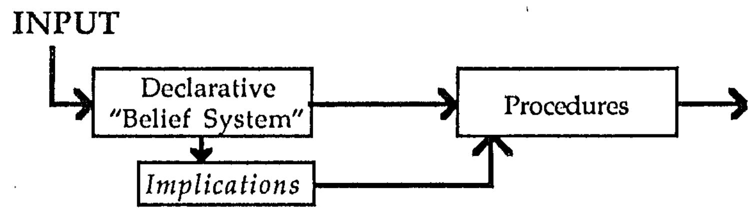


Figure 11

experiment 1a, subjects from the same population were able to produce significantly more principle codings when sorting mechanics problems.

The results of the first experiments led the author to reassess the direction of this thesis. The remaining part of this discussion surveys the literature relevant to the results and introduces a new working model to guide further research.

A New Model

The working model of figure 9 is not very consistent as the progression from novice to expert is left unstated. To express this, as well as the rigidity of novices' procedures a new model was devised (figure 11). *Implications* are the knowledge that is normally implicit within the problem solving process. This information describes the appropriateness and gives application guide-lines. The model is not intended to be exact, as the directional links with the *implications* are likely to be two-way. Instead it is a model that illustrates one of the possible expert - novice differences. That is, novices could have fewer *implication* rules than experts.

The past research on these aspects of implicit knowledge will now be described, followed by a summary of *implications*. The second-order features found by Chi et al. almost read like a check-list of appropriateness:

Is the "force too complicated"?
Is there a "before and after situation"?
Are there "given or well defined initial conditions"?
Is the problem a "determination of something at an instant in time"?
Is there an "elastic collision"?
Are "interacting objects" involved?

Although as far as the author is aware no one has suggested a purpose to the second-order features, their role in this context is quite clear. The second-order features are quite likely the criteria that experts use to determine the appropriateness of procedures. The search domain maybe narrowed by the principles that are uncovered and then further reduced by applying a check-list of appropriateness.

Reif and Heller (1981) present a detailed prescriptive model of physics in which *implications* are referred to as ancillary knowledge. Their ancillary knowledge includes "knowledge used to interpret concepts and relations". This involves defining a problem so that it is practical and unambiguous. They also include knowledge of algebra that enables equations to be rearranged or transformed (e.g. differentiation).

Reif and Heller's applicability conditions (appropriateness) are more confined and specify when it is legitimate to use particular concepts and principles. For example, motion principles are accompanied by an appropriateness rule that specifies that they can only be applied to a system when its motion is given relative to an inertial frame. Another type of ancillary knowledge specifies the kind of output that is acceptable. This important aspect allows the resulting output to be checked as well as providing information before the procedure is executed to determine whether the desired goals will be satisfied.

Finally Reif and Heller describe heuristic rules within which they include "advice about when particular knowledge might usefully be applied". In many circumstances there are several principles that can be applied to get the desired result, however some methods require more work or

Errors in specification of concept
- Gross confusions
  - Confusion with concept denoted by similar symbol
  - Confusion with concept describing different features of same situation
- Errors in specification rules
- Errors in applicability conditions

Errors in specification of values
- Errors in specifying ingredients
- Errors in possible values

Errors in specification of independent variables
- Omitted independent variables
- Wrong independent variables or properties thereof

figure 12

provide a more detailed output than is necessary. In terms of the *implications* model this aspect is considered a sub-group of rules of appropriateness.

For this thesis *implications* are limited to ancillary knowledge specific to a particular procedure. An incomplete set of ancillary knowledge could then explain the errors that are made by novices while solving a specific problem. It is foreseeable that a collection of procedures would have a set of common *implications* and further ancillary knowledge would exist to specify such things as the best interpretation and creation of an initial representation .

Problem solving errors is a research area in its own right. One characteristic that makes an expert is a well developed error and detection system. Reif (1985) lists the most common errors that can occur during problem solving (following as apriori analysis), figure 12.

The most notable error made by subjects in experiments 1a and 2 was to omit independent variables (e.g. the kinetic energy of m_2).

The application of certain principles or procedures can be seen to have a set of associated information pin-pointing the areas most likely to develop errors. Errors can then be avoided or detected and corrected. Rules of appropriateness for a procedure could include reference to the quality of the associated error knowledge. A procedure that has been used more often will likely have a better set of error pointers as well as a more automated form.

Other researchers have studied in depth the systematic errors ("bugs") made by subjects problem solving in highly constrained domains. Most of this research follows a similar approach to artificial intelligence with simulation programmes devised to emulate the errors of novices. One such well known study was run by Brown and Vanhehn (1980). After a detailed analysis of the errors children make performing place-value subtraction (figure 13), they formulated a "repair theory" with a set of principles to explain the incorrect or missing information and another set to describe the repairs that are made.

$$\begin{array}{r} 1245 \\ \underline{-\ 598} \\ 647 \end{array}$$

figure 13

In general the first set of principles involved the deletion of a rule and the execution of its sister rules instead. Little emphasis is given to ancillary knowledge, however without knowledge of what an error looks like, repair heuristics will never be established. The error detection is incorporated by specifying the limits of the input and output of a rule (operation). For example, the rule: "Decr -- Subtract one from the digit contained in the argument and writes the result back in the same cell". ("The input digit must be lager than zero"). Repair processes take over when a rule's operation exceeds its parameters. Extending this idea to physics and problem solving in general, a procedure can have input and output parameter limits. These could specify that particular ingredients or independent variables (figure 12) that are required as well as limits of singular or combined input information.

The "input digit must be larger than zero" is another example of an appropriateness rule. Ohlsson (1987) refers to such information as implied[1] (P implies Q) in contrast to productions (When conditions, Do action). Unfortunately Ohlsson perceives productions rather than implications as the "language of appropriateness". However, this difference is likely only due to contextual differences and understanding of what is appropriateness. To explain the role of implications and productions in more detail as the author of this thesis sees them, the previous Decr rule will be used.

IF conditions THEN Decr

IF Decr (executed)
THEN input digit must be larger than zero

The first line is a condition-action production that is neither true or false, while the second is an *implication* that has a true or false value based on the proposition "input digit must be larger than zero". If the proposition is false an error has occurred. Ohlsson assimilated *implications* with declarative knowledge and productions with procedural knowledge, going on to describe the interaction between the two. However for this thesis the concept of implications will be limited to those associated with a particular procedure.

Returning again to physics , part of the inability of subjects to solve the problems of experiment 1a can be described in the following way. Novices, due to their inexperience, have acquired only a few procedures that have a limited domain of application. However, when a problem is presented that has the same surface characteristics but is outside the solvable domain the erroneous procedure is activated. That is, novices have taken a specific procedure and generalized it to another problem without the ancillary knowledge indicating the valid variations and consequences (this is known as a "slip"). As mentioned this knowledge consists of operational limits of subsets of the procedure. With experience these *implications* are expanded and refined. Reif and Heller describing the problem solving process say "the most difficult decisions are those made to generate constraints (i.e. decisions about what principle to apply, to what system, at what time, and with what description)".

In summarizing *implications* a brief list of the information they can contain, follows: input requirements - specific variables and their format (e.g. flow diagrams), accepted values; output form and consequences - specific variables and their format (e.g. level of detail), accepted values, complexity of procedure (e.g. expenditure of time and mental resources).

The purpose of this section was to express the usefulness of ancillary knowledge and its coherence with the literature. To this end their relationship to declarative knowledge and productions was touched upon.

Naive Physics

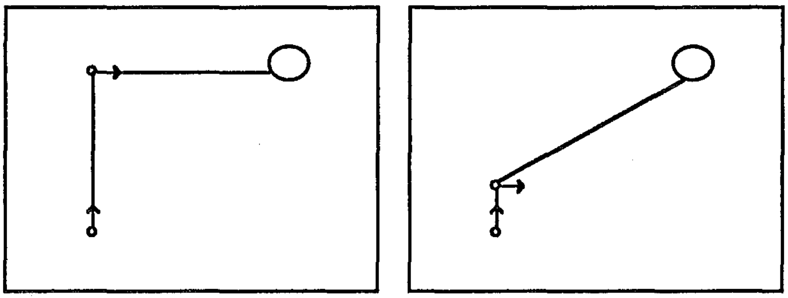
The experiments to be reported in Chapter 4 include a brief test of a naive physics misconception using a test called the coin problem. To put this test in context this section will give a brief overview of the research on naive physics.

The general conclusion of this area of research is that naive subjects have strong Aristotelian expectations about the world. These expectations are quite robust and continue to influence behaviour after several years of formal physics education.

The Aristotelian view under greatest study is the idea that a moving object must have a force acting on it, in the direction of the motion. DiSessa (1982) created a simple video game in which subjects had to guide a rocket to a target. They were able to control the timing of the rocket fire and the

---

[1] Ohlsson also uses the term implications as distinct from productions. This is the original source of *implications* as they are referred to in this thesis.



a b

figure 14

direction. DiSessa found that school children had definite non-Newtonian expectations. The most common expectation is shown in figure 14a (and the correct action in figure 14b). A first year university student, however, showed a combination of Newtonian and non-Newtonian strategies.

Clement (1982) had subjects answer the following coin problem (figure 15):

> A coin is tossed from point A straight up into the air and caught at point E. On the dot on your paper draw one or more arrows showing the direction of each force acting on the coin when it is at point B. (Draw longer arrows for longer forces).

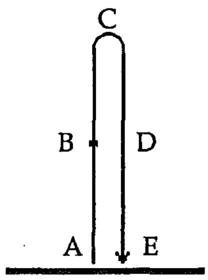


Figure 15

Nearly all the first year engineering students (equivalent to Canterburys first year physics) put an upward force acting at point B. An incorrect answer was given by 88% of the students. At the end of a year of engineering the percentage of incorrect answers dropped, but 75% were still wrong!

People do not just stick to the Aristotelian view of motion. McCloskey, Caramazza and Green (1980) found first year students exhibiting historically earlier impetus theories of motion. They had subjects predict the course of a ball spinning around on a string when the string broke. Thirty percent of the subjects thought the ball would take a curved path (figure 16).

The coin problem is included in the experiments of Chapter 4 to replicate the findings and to check the extent of the misconception with varying degrees of expertise.

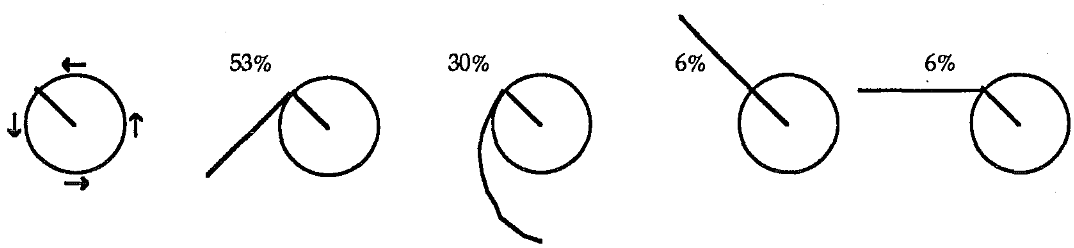


figure 16

CHAPTER FOUR

Research

Experiment 3a

Experiment 1a (involving the solving of a number of physics problems - P1, P2, P3, P4) is repeated here, however the novice subjects were third year physics students.

METHOD

Subjects

Twelve third year physics students (9 male and 3 female) volunteered from the 18 students approached in a laboratory class.

Materials

Problems 1 to 4 (P1, P2, P3, P4) of experiment 1a were used. However P4 was changed slightly to "What is the speed of m_2 when the man reaches the ground" (In this way the velocity of m_2 and therefore kinetic energy are hard to ignore in the application of the *conservation of energy).*

Procedure

Subjects in the control group were asked to solve P4 first. The instructions and procedure were identical to experiment 1a. The P1, P2, P3, P4 treatment was not applied due to the failure of subjects to solve P4 in the Control group.

RESULTS

None of the first four subjects in the Control group solved P4. One of the subjects applied the simple theorem $m_1 v_f^2 = m_2 gx$ for the *conservation of energy* (despite the question asking for the velocity of m_2). The other three students in the Control group used Newtonian physics incorrectly or combined it with differentiation to avoid stating kinematic relations (though they were in fact doing so). The results for experiment 3a and 3b displayed quite a distinct bimodal distribution. This troublesome arrangement occurred between the Honours school students and the third year students taking the same laboratory class. Honours students tend to be those for whom physics holds high intrinsic interest while many of the remainder were studying physics very much as a second choice because they had not gained entry to engineering school. In general the Honours students applied far more complex methods than was necessary, while the attempted solutions of the remainder were like those of the first year students of experiment 1a.

Discussion

Unfortunately little can be gleamed from the experiment. As was shown in experiment 1a students do have procedures that they can apply to the situation. However, as before, information relating to the appropriateness of these methods is either non existent or is not considered.

Experiment 3b

This experiment is a continuation of the subjects from experiment 3a (which failed when the subjects in the control group was unable to solve problem 4). The third year physics students sorted by *principle* and *similarity in solution* , as well as answering the coin problem used to test concepts of naive physics. It was expected that third year students would display the same sorting groups as novices. A further question was asked to probe the *implications* of principles available to the subjects.

METHOD

Subjects

Eleven third year physics students from experiment 3a took part in this experiment.

Materials

The sort cards from experiment 1b were applied. The coin problem was also given (see Appendix D).

The following *implications* question was asked:

> "Referring to Problem 17, would you expect the principle of *conservation of energy* to apply between the initial state for when the bullet and the block are separate entities, and the state when they are moving with a common velocity? Explain your answer." (Problem 17 is given in Appendix B)

Procedure

The sort task was the first performed by either *principle* or "similarity in solution" (five students in one group and six in the other), with the standard instructions applying. Following the recording of the sort results, subjects were asked to answer the *implications* question. Subjects were given no time restraints, however each interview was scheduled to take 30 minutes. Finally the students answered the coin problem.

All subjects performed the sort task and the coin problem. Seven subjects were asked to answer the *implications* question.

RESULTS

As mentioned in experiment 3a a bimodal distribution developed in the sort task results, due to a mixture of Honours and the Standard class in the laboratory. Figure 17 shows the subject numbers in each group.

	principle	solution
standard class	3	3
Honours school	3	2

Figure 17

All the P values (using the coding method described in experiment 1b) are presented in Appendix E. However figure 18 (over the page) shows the P values for the Honours and Standard class. Using a Wilcoxin Rank-Sum test, normal approximation and correction for ties as in experiment 1b, the Honours and Standard class difference is significant ($p<0.5$ for two-tail). The *principle* and solution group difference is not significant.

The Standard class - solution group produced group labels typical of novice students. One of the three subjects sorted as naive students do (See experiment 5). The Honours - solution group, however sorted like experts. For example:

> "frames of reference"
> "impulse over time - equivalent to force
> "springs - Hooks law and conservation of energy"
> "conservation of momentum"
> "conservation of angular momentum"
> "conservation of energy"
> "force - Newton`s 2nd law"

The coin question also produced a bimodal split.

Frequency of P values

	P values 0	0.25	0.33	0.57	0.71	0.75
Honours	0	1	1	0	2	1
Standard	3	2	0	1	0	0

$S_c = -22$
$z = -2.05$
$p = 0.04$ (two tail)

Figure 18

One of the six students from the Standard class answered the question correctly (i.e. excluding an upward force associated with velocity). Four of the five students from the Honours class answered the question correctly. The following descriptions were given for the imaginary upward force:

"momentum of coin"
"force from toss"
"force associated with the coin's velocity"

The *implications* question was answered in two ways. One focusing on the nature of the collision and the other on the conservation within a specified system. For example:

> "No, some of the energy of the two seperate systems would be lost to the final system in the impact. That is some of the energy of the two seperate entities would be lost in the collision being converted to sound and heat."

> "Can consider bullet, block and spring as an isolated system - no energy passed onto or out of system, only among components. Conservation of energy is a univesal law; isolating the system only allows easier computation."

Discussion

As the subjects of this experiment have more experience than those of experiment 1a, a slight increase in *principles* was expected. However the large jump between the Standard and Honours class suggests aptitude plays an important role in qualitative changes in knowledge structures.

General comments about the *implications* question are hard to make because of the varying answers supplied. Each answer was flawed due to a lack of certain contextual details. That is, information that specifies the instance and assumptions that the answer refers to. The answers given were appropriate if the assumptions made were explicated. For example the second answer (above) is correct if the system includes the heat loss, etc. from the collision, or refers to an ideal textbook problem with mechanical energy conserved. However considering the information given overall, a good understanding of the application of the *conservation of energy* is shown, with no notable information missing with all subjects. The system and its before and after elements are discussed as well as the idealized situation of no energy lost in the collision.

Experiment 4

To comment any further on the results and the inability of students to solve the questions, the responses of two experts were examined. Up until this point it had been assumed (based on advice from members of the academic staff in Physics and Mechanical Engineering) that students as well as experts could solve the problems given. The results of experiment 1 - 3 cast doubt on the ability of students to solve the problems therefore it is necessary to test the assumption that experts can solve problem 4 !

The first expert did not consider the velocity of m_2 in problem 4 (the same mistake novices made), even though this was what was requested. Problem 4 was made more explicit for the second expert.

METHOD

Subjects

Two mechanical engineering doctoral students supervised by Professor MaCallion willingly agreed to participate. These two students were judged to be experts by Professor McCallion having spent their previous years involved with mechanics requiring the application of *conservation of energy* and Newtonian physics.

Material

From experiment 3 the following problems were used: problem 4, the *solution* sort task, the *implications* question of experiment 3b and the coin problem.

Problem 4 was presented differently to each subject:

1. What is the speed of m_2 when the man reaches the ground? (Derive a formula for the speed without using kinematic relations e.g. $v^2 = u^2 + 2as$).

2 What is the speed of the <u>block</u> m_2 when the man reaches the ground? (Derive a formula for the speed <u>without</u> using the acceleration of the masses).

Procedure

The subjects were given the instructions and timed while they answered problem 4. If the answer was incorrect they were asked to try again. On completion of problem 4 the two experts performed the sort task by "similarity in solution".

Next the *implications* question was administered and finally the coin problem.

RESULTS

Problem 4 provided surprising responses from both subjects. Subject 1 did not include the kinetic energy term for m_2 in his calculations and when he failed to solve the problem, he reverted to using Newton's laws of motion (though not successfully). Later questioned on the explicit reference to the speed of m_2, the subject reported he misread the question as asking for the speed of m_1. Subject 2 chose to ignore the "<u>without</u> using the acceleration of the masses" clause and applied Newtonian physics. He also incorrectly equated the acceleration of the masses with the acceleration due to gravity and therefore failed to solve the problem. His explanation for using the Newtonian approach was that he "avoids the conservation of energy like the plague", because he is far less familiar with it.

The two subjects performed differently for the sort task. Subject 1 sorted the cards into four *principle* groups:

"Newtonian laws, F=ma"
"conservation of energy or energy principles"
"a combination of the above two"
"conservation of angular momentum"

Subject 2 sorted as follows:

"projectile motion"
"friction and movement"

"angular momentum and angular kinetic energy"
"impulse and momentum - collisions"
"non-dynamic forces - involve just forces - solved with force diagrams"
"gravity and Conservation of Energy methods"

A check of the contents of the groups for subject 2 indicated that three could be coded as *principle* labels. This is based on the group name hinting at a *principle* grouping plus the majority of the cards within that group requiring the *principle* in the solution.

The *implications* question also provided diverse answers. Subject 1 mentioned the objects comprising the system, the initial kinetic energy of the bullet and none for the block, and an inelastic collision. Subject 2 stated that the principle could be applied as the loss of mechanical energy from the collision was negligible due to the size of the masses and speed involved.

The coin problem was answered correctly by Subject 2. The common misconception of the existence of an upward force, however, was strongly held by Subject 1 (a few minutes were required to convince the subject that this was so).

Discussion

The failure of the two subjects to solve problem 4 questions the whole methodology of this research on the *conservation of energy* and pulley problems. This may have been an unfortunate choice. However the two experts of this study may be novices in their ability to apply the *conservation of energy* to pulleys. Subject 2 expressed his distaste of energy theorems and indicated their secondary role in his problem solving repertoire. The simple lack of practice of the principle *conservation of energy* and therefore a reduced set of *implications* may explain his results.

The lack of consideration for the speed of the block m_2 as the man touches the ground may, in fact, be a specialized misinterpretation that is confined to this very problem. It could be that only well developed trouble-shooting / error checking *implications* of experts will enable the experts to avoid this pitfall.

Subject 1 did write several of the implications expected of experts. The low number of *implications* mentioned by subject 1 and the fact that Subject 2 mentioned none, may reflect the inadequacy of the question rather than low *implicational* knowledge. This is supported by the unrecorded verbal protocol of a physics lecturer as he analysed the question. He said that due to the large loss of energy in the collision the application of the *conservation of energy* would be far less appropriate than applying the *conservation of momentum*.

The *implications* knowledge appears to be difficult to tie down. The only effective method of determining such knowledge may be through analysing explicit protocols from experts as they problem solve (as in part Chi et al. did).

After four years of university physics / mechanical engineering Subject 1 still displayed an Aristotelian perception of the movement of the coin. This supports Clement's (1982) finding that people in general have an Aristotelian understanding of the world. And this "belief system" can remain throughout formal physics / engineering education. Such results are also supported by Professor MaCallion's own personal observations of engineering textbooks and academic papers published with fundamentally basic flaws.

Experiment 5

The failure of the two subjects to solve problem 4 in experiment 4 and the low *principle* score in the sorting task by one of them, questions the definition of an expert. Five physics lecturers (experts) were required to sort by "similarity in solution" and then by *principle..* The original plan was to select six lecturers specializing in mechanics and six that did not. However this was abandoned as most lecturers were involved with tutoring mechanics and the lecturers that agreed to help were more likely to be confident at solving mechanics problems. *Implications* questions were also presented. These questions are combined with the results of experiment 6.

A second group of naive subjects (with no university physics experience), sorted by "similarity in solution" to provide a baseline for discussing the novice sort results.

METHOD

Subjects

Six physics lecturers agreed to participate and served as the expert group. One lecturer withdrew part way through the sort task prefering not to continue, leaving five subjects.

Five psychology masters students and one layperson served as the naive physics students. None had completed a year of university physics and none had studied physics (e.g. school physics) in the past five years.

Materials

The standard sort cards from the previous experiments were used. An *implications* question reported in experiment 6 was also used.

Procedure

The experts were first asked to sort the cards designed for experiment 1b by "similarities in solution". Following the recording of these groups and group labels they were asked to re-group the divisions they had made so that the cards were sorted by "fundamental principles or laws of physics". Any changes were recorded.

Following the usual demographic questions the experts were asked to fill in the *implications* questions that are reported in experiment 6.

Members of the naive group were asked to sort the cards into groups accoring to "similarities in solution".

RESULTS

Due to the low subject numbers for experts, all five sorts are presented here:

1. "conservation of momentum and energy"
   "conservation of angular momentum and energy"
   "conservation of angular momentum"
   "kinetic and potential energy"
   "kinetic energy and work against friction"
   "impulsive force"
   "F=ma"
   "F=ma and Hook's law"
   "Hook's law"
   "straight calculation of potential energy"
   "accelerated reference frames"
   "equilibrium under torque"
   "insufficent information"
2. "F=ma Newton's first law"
   "conservation of momentum and energy"
   "conservation of energy"
   "torque, total torque = zero"
   "conservation of angular momentum"
   "energy"
   "force constant of spring"
   "Newton's second law"
3. "F=ma Newton's first law"
   "forces in statics"
   "conservation of linear momentum"
   "conservation of angular momentum"
   "friction (kinetic)"
   "energy in a spring - relaease of kinetic energy"

"potential and kinetic energy interchange - conservation of energy"
"conservation of energy and linear momentum"

4. "collisions with conservation of angular momentum and energy"
   "collisions - linear momentum"
   "conservation of energy + spring constants + friction"
   "simple"
   "simple lever"
   "explanations -understanding of principles"
   "forces (can't use conservation of energy) - simple dynamics"

5. "one dimensional force-momentum (e.g.F=ma)"
   "vector force-momentum"
   "impact - using conservation laws"
   "energy"
   "conservation of angular momentum"

All five experts were happy with the groups previously created when asked to re-sort according to *principle*. Several of the experts verbally confirmed that they had set forth to sort by *principles* in the first instance.

The naive subjects produced groups that had no deep structure rating on the scales developed (e.g. #Prin). However their group labels were distinctly different from those of novice physics subjects. All naive subjects chose group labels from physics terms that were explicitly written in the problems question. For example, " ... What is the velocity of the mass m_2?" is placed in a group labeled "velocity". A typical response was as follows:

"angular velocity"
"speed"
"acceleration"
"distance"
"?"
"force"
"direction"
"tension"
"potential energy"
"kinetic energy"

(A ? indicates a group of problems that the subject could not classify).

Discussion

The responses of the naive subjects were distinctly different from the novices of experiment 1b (first year physics students). Therefore the initial formal years of education in physics have had a notable effect on the structure of physics knowledge.

The experts conformed to the experimenter's expectations and the findings of Chi et al. The large differences between experts previously reported, were also found here. These differneces exist while maintaining the high proportion of principle categories. Experts differ in the problems that go in the same principle categories as well as the extent to which surface features are mentioned.

Experiment 6

Experts and novices were questioned on their preference for using the conservation of energy or Newtonian physics in situations where the choice existed. This was done to test the assumption that experts use effective appropriateness rules (*implications*) when problem solving and novices have little understanding of such things.

METHOD

Subjects

Thirteen physics lecturers and fourteen second year physics students replied to the questionaire distributed to them.

Materials

The questions in figure 20a were presented to experts as well as the usual demographic information (i.e. age, sex, experts - what proportion of your time is spent solving simple or complex mechanics problems). Novices were presented with a slightly different form to give a better gauge of their understanding of multiple solution methods (Figure 20b).

Procedure

The questions were in a questionaire form and were placed in the lecturers mail pigeon holes to be completed in their own time. The forms were handed to novices during a lecture by the lecturer and they handed them in at the next laboratory class.

RESULTS

Responses were received from thirteen lecturers. One was discarded as the lecturer did not answer the questions as directed. Of the remaining twelve, two had no perference and the other ten preferred to use the *conservation of energy*. These experts said that the energy method, if applicable, deals with "end results", avoiding any possibe calculus intergration. It is simpler (more direct), gives comceptual clarity, is quicker, easier and is more likely to be error-free.

The novice class was split into Standard and Honours, with five of the twenty four Standard class returning their forms and nine of the twelve Honours class returning theirs. All but one of the Honours class said that there were many occasions where both methods can be applied. The novices' preferences are shown in figure 21 (over the page).

The novices' answers are not as easily summarized due to the larger individual differences. Seven novices stated specific physical conditions that determined when one method was better. For

There are many occasions when either Newtonian equations of motion or the Conservation of Energy can be applied to determine unknown values. In these circumstances which do you prefer to use ?

no preference ( ) Newtonian equations of motion ( ) Conservation of Energy ( )

If you need to qualify when you prefer to use one approach before another, please do:

Why do you prefer one approach over the other ?

Figure 20a

Are there many occasions when either Newton's laws of motion or the Conservation of Energy can be applied to a problem to determine unknown values?
Yes / No (circle)

Briefly explain your answer:

In the circumstances where both can be applied which do you prefer to use ?
no preference ( ) Newton's laws of motion ( ) Conservation of Energy ( )

When do you prefer to use one approach before another?

Why do you prefer one approach over the other ?

Figure 20b

	no perference	Newton's laws	*conservation of energy*
standard	2	0	3
Honours	3	2	4

Figure 21

example, "If forces, masses and accelerations are known or most easily meassured, I use Newton's laws". Six mentioned that the conservation of energy method was easier. One mentioned that less mistakes were likely to be made with this method. Two mentioned Newton's laws of motion provided more detail.

Discussion

The high level of appropriateness knowledge displayed in the responses of the novices was unexpected and runs counter to the view perpetrated earlier. However due to the low level of experimental control the questionaires returned were likely to have been a selective group that felt comfortable answering the questions. Further to this, it is not known whether the students were answering what they thought they **should** do, rather than what they **would** do.

Regardless of this the novices showed that they have acquired some form of *implicational* knowledge. The responses of the seven novices that provided specific physical conditions parallels the emphasis on surface features by novices, when sorting.

CHAPTER FIVE

General Discussion

The introduction to this thesis proposed a method to determine the generality of the surface - deep sort structure difference reported by Chi, Feltovich and Glaser (1981), and to test the effect of problem order. Unfortunately the question order which tested hypothesis formation, had no effect and the problems could not be solved.

At that point the research was redirected to discover the reasons for the low problem solving ability of subjects, focusing on the ancillary knowledge. Further to this the sort task instructions and level of expertise were varied to test the generality of the sort findings. This discussion will first concentrate on the results from these two areas (the sort task, and ancillary knowledge), followed by a section on production systems, to help place the results in the context of other research. Finally, implications for future research and education will be considered.

The Sort Task

One of the interesting findings of experiment 5 was the difference in sort behaviour of novice and naive subjects. A distinct difference between novice (first year physics) and naive (psychology masters) students was evident even though both groups produced labels that originated from the surface content of the problems. It would be interesting to ask whether the structure shown by the novices is domain specific or if the progression from naive to novice sorting can be found in domains other than physics. This is an avenue for future research.

The discovery of a naive level of sorting has implications for the way sorting features are defined. Chi et al. have referred to second order features as being derived from certain surface features. However in light of a naive sorting structure it is necessary to distinguish between the novice's and naive subject's surface features. In keeping with Holyoak and Koh's (1987) terminology the naive subject's direct excerpts (first order features) can be called superficial features while the novices sort features remain surface features. Second order features can then be said to be derived from certain superficial features. Silver (1979) in his study of mathematics problem solving refers to superficial features as the pseudostructure dimension, when problems are sorted according to a measurable quantity such as weight or time.

Since first year physics students of experiment 1b can sort by *principle* (28.6% of the cards) when asked to, and first year engineering students can distinguish between *principles* and structural properties (experiment 2), it seems that the students have some of the declarative knowledge necessary for solving mechanics problems. However, when they have to apply this knowledge they are hampered by a lack of appropriate procedural knowledge.

The bimodal split of experiment 3b shows striking differences between the Honours and Standard class. A difference due to the selection criteria of the Honours class is expected. However, such a large difference leads one to several possible conclusions:

1. The deep - surface structural difference is not due to maturation (i.e. the Standard and Honours class students are the same age), confirming one of Schoenfeld and Herrmann's (1982) conclusions.
2. The change from sorting by surface features to *principles* can be rapid (also demonstrated by Schoenfeld and Herrmann).
3. The speed of progression is likely to be dependent on aptitude as well as the amount of knowledge acquired through formal instruction (though these two are related).

Experiment 5 also confirmed Chi et al.'s findings by showing that the experts (lecturers) sorted by fundamental principles or laws of physics when asked to sort by solution. Also variations by experts within this *principle* level were evident, with several different groupings produced.

The results of the two Ph.D students (experiment 4) suggest that attaining a high level in the education system is no guarantee of expertise. This is not surprising considering a time gap of 10 to 30 years exists between the Ph.D students and the lecturers examined, which reinforces the notion of the effectiveness of practice over a long period of time. A further point is that lecturers are a more selective population than Ph.D students. Therefore the difference may be due to aptitude.

Ancillary Knowledge

The use of *implication* questions in the early stages of development (experiments 3 and 4 of Chapter 4) was not particularly helpful. One important aspect noticed, however, was the lack of contextual information indicating when certain procedural knowledge is appropriate. As this lack of contextual information could have been due to the nature of the question, the appropriateness of specific problem solving methods was tackled directly in experiment 6 . Unfortunately this too had its shortcomings. The questionnaire return rate was quite low. Therefore, it is possible that the results only reflect the knowledge of the more able students.

The *implication* questions were not devised to test hypotheses, but were intended to explore what *implication* knowledge had been acquired by the subjects. Experiment 6 showed that some of the students did have forms of *implication* knowledge. However, the role that this knowledge plays in the problem solving of novices remains unclear. One student demonstrated he was aware of the simplicity of the conservation of energy method, but stated that if he was in an exam situation he would probably use the method that first came to mind.

Implication knowledge is a useful concept in so far as it focuses research on a changing aspect of problem solving knowledge. The effects of concentrating teaching time on the acquisition of such knowledge needs study.

Production Systems

As this thesis has dealt with two different aspects of the novice - expert continuum (sorting and *implications*) the discussion will be more coherent if the results are described in terms of one encompassing theory. Anderson's (1987) ACT* theory serves this purpose.

Anderson has shown that procedures appear to be acquired in a very inflexible form. Singley and Anderson (1986, cited in Anderson, 1987) had subjects learn three computer text editing procedures in a specific order. By examining the features the editors shared and timing the performance on each, the researchers found that there was a strong positive transfer from one to the other only for the two editors that were composed of common productions. For example a production common to both maybe:

IF the goal is to type the command
THEN set as subgoals
1. To type the command name
2. To type the arguments

McKendree and Anderson, in press (cited in Anderson, 1987) have further tested the ability of subjects to transfer productions from task to task using the learning of LISP functions. Subjects were given several LISP functions to evaluate. Later they were given the results and arguments and asked to generate the function that would transform the arguments to the results. In terms of productions there was no exact overlap between evaluation and generation. For example the function LIST would have the evaluation form:

IF goal is to evaluate (LIST X Y)
and A is the value of X
and B is the value of Y
THEN (A B) is the value

and generation form:

IF goal is to code (A B)
THEN use the function LIST and set as subgoals
1. To code A
2. To code B

The required declarative content was the same but the procedures required to answer the questions were different. The researchers found little transfer between the two tasks.

Anderson's research has shown almost complete transfer between tasks if both the knowledge and the procedures are the same. However a change in procedures produces almost no transfer at all. A rigid procedure progresses to a more flexible form through two types of generalization. The conditional parts of more than one production are recognized as similar and a common production created. Alternatively, constants within a production are changed to variables. The latter is similar to the flexibility shown in the *implications* model of Chapter 3.

If the sort task is described as a procedure that structures the knowledge, the novice-expert distinction can then be described in terms of differing productions between experts and novices. At this stage the following important question must be asked: Is the sort task measuring problem solving? For the sort task to be measuring a meaningful unit of problem solving the sort task would need to share productions with the procedures of actual problem solving. These productions are presumably going to be involved with defining and creating (discriminating and categorizing) the initial representation.

The interesting point with this approach is that it is not the declarative knowledge that is heavily structured but the procedural knowledge that enforces an apparent structure through the form of the productions. For example, a novice:

IF problem involves inclined plane
THEN find angle of incline with horizontal

and an expert:

IF before and after conditions known
THEN apply conservation laws

Productions of these forms may be called (executed) as an early part of the sorting procedure, thereby determining the category requirements.

This interpretation of the results has widespread effects. The finding of experiment 1b, that novices can interpret 28.6% of the sort problems in terms of the *principles* could previously have been taken as a measure of subjects' declarative knowledge. However it seems they may have the declarative knowledge but lack the procedural form necessary for problem solving that is activated by the more specific *principle* sort task. This is then a statement about procedural knowledge and not about the structure of declarative knowledge.

Chapter 3 described rules of appropriateness (*implications*) as having the form of true or false propositions associated with a production. If the proposition is true, the procedure is appropriate. Some researchers do not identify ancillary knowledge as such. Anderson (1987) writes about his declarative-procedural distinction with procedures consisting of only standard productions. *Implication* information is probably an integral part of the production, simply in the form of satisfying the conditional section:

IF a quick method is required
with little intermediate (revealing) knowledge produced
and Condition3 ...
THEN apply conservation of energy method

Suggestions for Future Research

The research presented here could progress in several interesting and potentially fruitful directions. This thesis has shown that there is an apparent transition for novices, between classifying by surface features and starting to use *principles*. This transition could possibly be used as a measure of the effectiveness of teaching ancillary knowledge (see Educational Implications). Another interesting aspect is whether the sort groupings created by novices are determined by the structure of the physics course. This can be tested by varying the educational approach or, more easily, by examining the naive and novice sort groups from other domains (e.g. mathematics).

The large Standard/Honours class differences of experiment 3b are likely due to aptitude and possibly the different course structure for the two groups. This could be explored further by dividing first year physics students into two groups according to their high school bursary marks and checking the sort groups.

The usefulness of the *principle* sort group could be examined by looking at the transfer of procedures from, practicing to sort problems into *principle* groups, across to problem solving.

Probably the most effective direction for future research is through the education system. By making curriculum changes based on research findings, the validity of the various views of ancillary knowledge could be checked.

Educational Implications

There have been many papers written on the educational implications of problem solving research. This section will deal only with implications directly related to this research[1].

[1]Suggested readings (other than mentioned in this section):

Tuma, D. T. & Reif, F. (Eds.) (1980) *Problem Solving and Education: Issues in teaching and research* Wiley & Sons, NY.

One of the major emphases so far in this thesis has been on the appropriateness of procedures. Heller and Greeno (1978) state that "not only does the higher skill individual have available a broader range of schemata, but solution procedures are associated with the schemata and the procedures are ordered according to their usefulness with respect to different problem types and situations". (p138) From this, the question must be asked: Can the appropriateness of procedures be taught effectively? That is, is it more effective to spend time on the appropriateness of procedures than on other aspects of a course (e.g. a physics course). Several researchers are advocating the teaching of ancillary knowledge. However it is only recently that its effectiveness has been tested in the teaching field (e.g. Labudde, Reif and Quinn, in press, cited in Reif, 1987). This is likely to be tested further as Anderson (1987b) has produced a whole paper promoting the usefulness of pedagogical research.

Reif and Heller (1982) suggest that all aspects of a physics course can be more effectively structured as well as explicitly teaching components that are usually left implicit. For example, teaching students "separately how to generate basic and theoretical descriptions of problems; ... how to search for a problem solution by decomposing the problem systematically and exploring the relevant decisions; and how to assess the merits of the resulting solution". They also stress that students be taught to structure their knowledge hierarchically and that this knowledge is accompanied by application guide-lines.

Reif (1987) suggests that students be given various typical and error-prone problems so they can "compile a repertoire of knowledge about special cases and common errors".

Larkin (1981) puts the process of learning physics neatly into perspective: "Physics has many so-called 'fundamental' principles that state in

Reif, F. (1987) Instructional design, cognition, and technology: Applications of the teaching of scientific concepts. *Journal of research in science teaching,* 24,309-324.

simple form broadly applicable knowledge". However to be able to use such principles the student needs to have acquired a vast array of ancillary knowledge. Larkin gives the simple principle of conservation of energy as an example: the principle "requires being able to recognize many very different kinds of energy (kinetic, potential, rotational, microscopic (sic), etc.) and to relate them appropriately". (p333)

All of these suggestions are consistent with the findings of this thesis. That is, students appear to lack effective ancillary knowledge or the ability to apply it. This is especially so of the lack of awareness of error-prone areas in a problem solving procedure.

The concept of naive physics misconceptions has been more of a side issue here than a focus of discussion. As a concluding note, McCloskey, Caramazza and Green (1981), and Clement (1982) decided that for effective teaching a student's preconceptions need to be taken into account. The correct aspects of preconceptions can be built on by having students describe and become aware of their own preconceptions. These can then be compared with other theories and conflicting situations. Without this a student is likely to distort what they are taught to fit their present misconceptions[1].

Conclusion

The sort task results of Chi, Feltovich and Glaser (1981) have been replicated by this thesis. A third distinct level of expertise was found for subjects with no physics experience using the sort task. Overall the results here have shown that novice subjects (1st, 2nd, and 3rd year university physics or engineering) have the relevant declarative knowledge but lack well-worn procedures necessary for problem solving. Likewise they have the *implication* knowledge necessary but this may not be in an applied/procedural form. The sort task results appear to reveal the procedural knowledge associated with problem solving, rather than the declarative knowledge. The sort task may share productions with an early stage of problem solving.

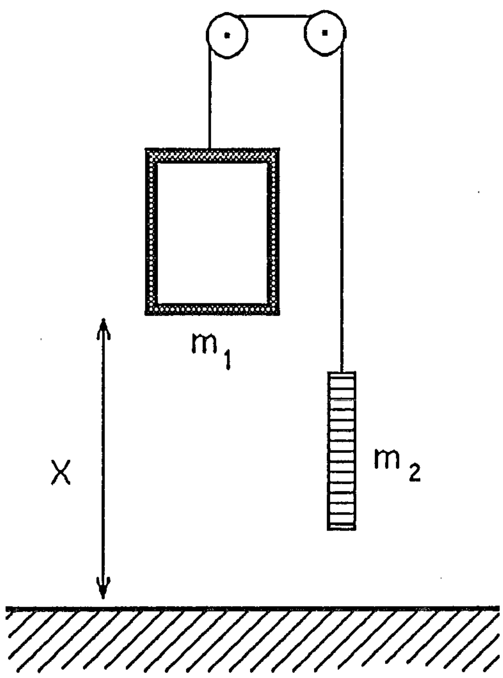
[1]It appears that it could be more appropriate to discuss naive physics preconceptions in terms of theories of communication rather than as unsuccessful problem solving.

Appendix A

Initial Instructions

The following mechanics problems have been adopted from the physics text book used by PHYS101. One question must be completed before moving onto the next question. I will be recording the time required to complete each question, so please tell me when you believe you have finished each one. Give all working out in full as you would in an examination, plus any relations / formula you are contemplating but may not use. Please write in order down the page. Are there any parts of these instructions that need explaining ?

For the following problems the acceleration of gravity = 9.8 m/s^2.



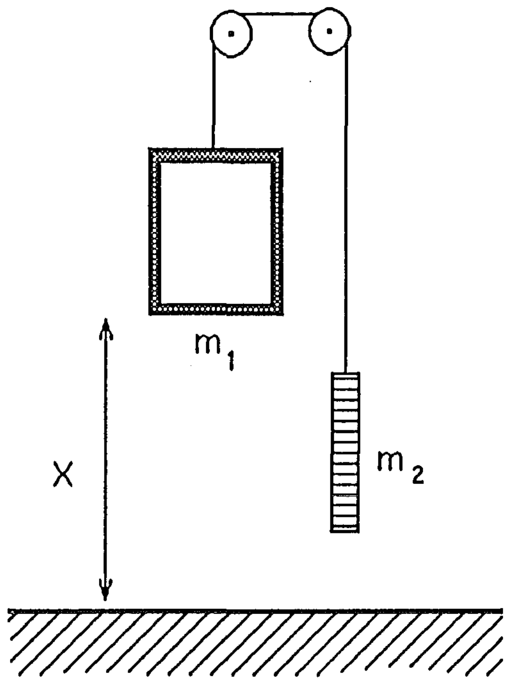
Problem 1

An empty elevator is an unknown height above the ground and consists of an elevator cage of mass m_1 (1200kg) connected by a cable, running over a pair of pulleys, to a counterweight m_2 (1300kg). Neglect the mass and friction of the cable and pulleys. What are the tensions in the cable if the pulleys are locked (by means of a brake) so that the elevator remains stationary ?

Problem 1'

What are the tensions in the cable if the pulleys are locked (by means of a brake) so that the elevator remains stationary ?

Given that an empty elevator is an unknown height above the ground and consists of an elevator cage of mass m_1 (1200kg) connected by a cable, running over a pair of pulleys, to a counterweight m_2 (1300kg). Neglect the mass and friction of the cable and pulleys.



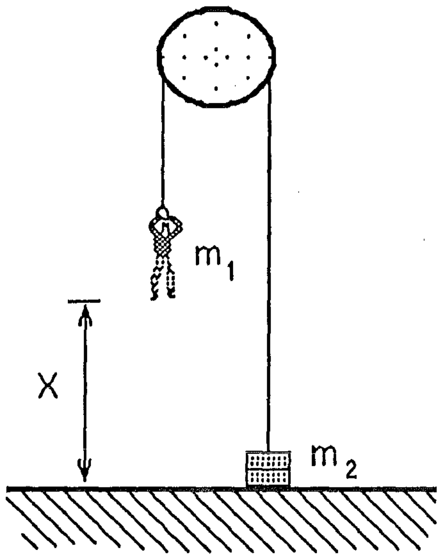
Problem 2

An empty elevator is an unknown height above the ground and consists of an elevator cage of mass m_1 (1200kg) connected by a cable, running over a pair of pulleys, to a counterweight m_2 (1300kg). Neglect the mass and friction of the cable and pulleys. What is the upward acceleration of the elevator cage if the pulleys are permitted to run freely? (Derive a formula for the acceleration and evaluate it).

Problem 2'

What is the upward acceleration of the elevator cage if the pulleys are permitted to run freely? (Derive a formula for the acceleration and solve).

Given that an empty elevator is an unknown height above the ground and consists of an elevator cage of mass m_1 (1200kg) connected by a cable, running over a pair of pulleys, to a counterweight m_2 (1300kg). Neglect the mass and friction of the cable and pulleys.

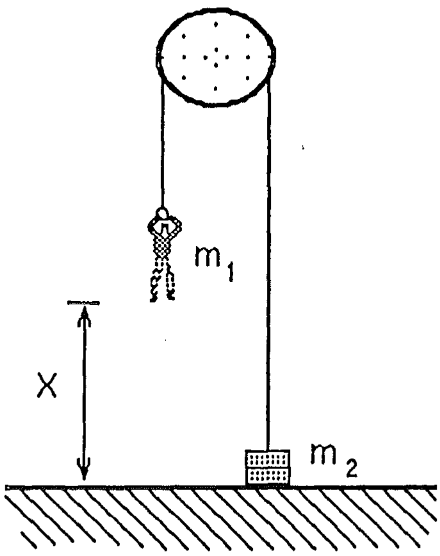


Problem 3

A man of mass m_1 (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass m_2 (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. What is the tension in the rope ? (Derive a formula for the tension and evaluate it).

Problem 3'

What is the tension in the rope ? (Derive a formula for the tension and solve). Given a man of mass m_1 (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass m_2 (70kg),which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley.



Problem 4

A man of mass m_1 (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass m_2 (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. How much energy does the mass m_2 gain by the time the man lands on the ground? (Derive a formula for the energy and evaluate it).

Problem 4'

How much energy does the mass m_2 gain by the time the man lands on the ground? (Derive a formula for the energy and evaluate it).

Given that a man of mass m_1 (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass m_2 (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley.

Problem 4a

A man of mass $\mathbf{m_1}$ (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass $\mathbf{m_2}$ (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. How much kinetic energy does the mass $\mathbf{m_2}$ gain by the time the man lands on the ground? (Derive a formula for the energy and evaluate it).

Problem 4b

A man of mass $\mathbf{m_1}$ (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass $\mathbf{m_2}$ (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. With what speed does the man hit on the ground? (Derive a formula for the speed **without using kinematic relations e.g. $v^2 = u^2 + 2as$**, and evaluate it).

Problem 4c

A man of mass $\mathbf{m_1}$ (80kg) lowers himself to the ground from a height x (5 metres) by holding onto a rope passed over a pulley and attached to a block of mass $\mathbf{m_2}$ (70kg), which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley.

a) What is the change in potential energy in the system? (a gain or loss?)

b) What is the total kinetic energy of the system as the man reachs the ground?

c) How much kinetic energy does the mass $\mathbf{m_2}$ gain by the time the man lands on the ground?

Problem 1

...What are the tensions in the cable if the pulleys are locked (by means of a brake) so that the elevator remains stationary ?

Answer 1:

If the pulleys are locked, the tension in the cable on either side must match the weight hanging from that side.

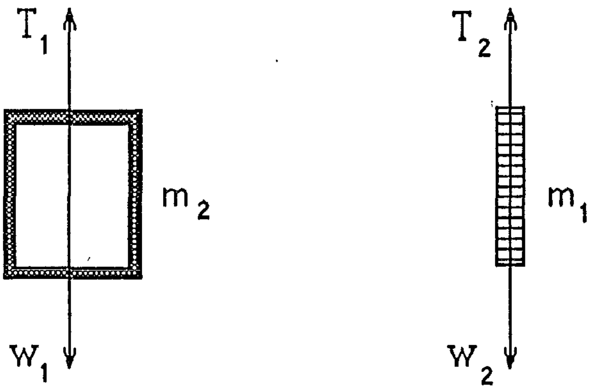
Thus:

$T_1 = m_1 g = 12000N$ $\qquad$ $T_2 = m_2 g = 13000N$

Problem 2

... What is the upward acceleration of the elevator cage if the pulleys are permitted to run freely? (Derive a formula for the acceleration and evaluate it).

Answer 2:



$T_1 - w_1 = m_1 a_1$ $\quad$ (w_1 = weight of mass m_1)
$T_2 - w_2 = m_2 a_2$

As m_1 & m_2 are tied together $a_1 = -a_2$.

$$T - m_1 g - (T - m_2 g) = m_1 a_1 - (-m_2 a_1)$$

T cancels out :

$$- m_1 g + m_2 g = m_1 a_1 + m_2 a_1$$

So:

$$a_1 = \frac{m_2 - m_1}{m_1 + m_2} g \quad ==> \quad a_1 = ... \; m / s^2 .$$

Problem 3

...What is the tension in the rope ? (Derive a formula for the tension and evaluate it).

Answer 3 :
With

$$a_1 = \frac{m_2 - m_1}{m_1 + m_2} g \qquad \text{and } T - m_1 g = m_1 a_1$$

$$T = m_1 \frac{m_2 - m_1}{m_1 + m_2} + m_1 g$$

$$T = \frac{2gm_1m_2}{m_1 + m_2}$$

$$T = \;....\; \text{N.}$$

Problem 4

...How much energy does the mass $\mathbf{m_2}$ gain by the time the man lands on the ground? (Derive a formula for the energy and evaluate it).

Answer 4 :

The initial energy is equal to the final energy, so:

$$\frac{1}{2}(m_1+m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1+m_2)v_f^2 + m_2gx$$

$v_i = 0$ therefore

$$v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1+m_2)}$$

Gain in energy of mass $\mathbf{m_2}$ when the man lands is:

$$= \frac{1}{2}m_2v_f^2 + m_2gx$$

Substitute the equation for v_f^2 into this equation . . .

Problem 4a

...How much kinetic energy does the mass $\mathbf{m_2}$ gain by the time the man lands on the ground? (Derive a formula for the energy and evaluate it).

Answer 4a :

The initial energy is equal to the final energy, so:

$$\frac{1}{2}(m_1+m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1+m_2)v_f^2 + m_2gx$$

$v_i = 0$ therefore

$$v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1+m_2)}$$

so K.E. of $\mathbf{m_2} = \frac{1}{2} m_2v_f^2$

$= ......$ J

Problem 4b

...With what speed does the man hit on the ground? (Derive a formula for the speed **without using kinematic relations** e.g. $v^2 = u^2 + 2as$, and evaluate it).

Answer 4b :

$$\frac{1}{2}(m_1+m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1+m_2)v_f^2 + m_2gx$$

$v_i = 0$ therefore $$v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1+m_2)}$$

Problem 4c

Answer 4c :

a) What is the change in potential energy in the system ? (a gain or loss ?)

There is a loss of P.E. $= (m_2-m_1)gx$

$= \dots\ J$

b) What is the total kinetic energy of the system as the man reachs the ground ?

As there is no loss of energy in the system due to friction...

the loss in P.E. must now be a gain in K.E. (Conservation of Energy)

K.E. $= (m_2-m_1)gx$

$= \dots\ J$

c) How much kinetic energy does the mass m_2 gain by the time the man lands on the ground ?

This can be calculated by looking at the mass proportions ...

$$\text{K.E. of } m_2 = \frac{m_2}{m_1+m_2}(m_2-m_1)gx$$

or using Conservation of Energy...

The initial energy is equal to the final energy, so:

$$\frac{1}{2}(m_1+m_2)v_i^2 + m_1gx = \frac{1}{2}(m_1+m_2)v_f^2 + m_2gx$$

$v_i = 0$ therefore $$v_f^2 = \frac{2(m_1 - m_2)gx}{(m_1+m_2)}$$

$$\text{so K.E. of } m_2 = \frac{1}{2} m_2v_f^2$$

$$= \dots\dots\ J$$

Appendix B

The Sort Problems

1

A frame of mass 200g, when suspended from a certain coiled spring, is found to stretch the spring 10cm. A lump of putty of mass 200g is dropped from rest onto the frame from a height of 30cm. Find the maximum distance the frame moves downward?

2

A large disk has a mass 2kg, radius 0.2m, and initial angular velocity 50 rad/s, and a small disk has a mass 4kg, radius 0.1m and initial angular velocity 200 rad/s. Find the common final angular velocity after the disks are pushed together.

3

Two masses m_1= 1.5 kg and m_2= 3.0 kg are connected by a thin string running over a massless pulley. The mass m_2 hangs from a string; the mass m_1 slides on a 35^o ramp with a coefficent of kinetic friction = 0.40 . What is the change in potential energy in the system when m_2 has dropped one metre?

4

Two automobiles both of 1200kg and both travelling at 30km/h collide on a frictionless icy road. They were initially moving on parallel paths in opposite directions, with a centre-centre distance of 0.1m. In the collision the automobiles lock together, forming a single body of wreckage; the moment of inertia of this body about the centre of mass is $2.5 \times 10^3 kgm^2$. Calculate the angular velocity of the wreckage?

5

A block on an inclined icy slope reaches the bottom with a speed of 19.8 m/s where it moves over a very rough surface, quickly coming to rest. If the kinetic friction is 0.5, how far does the block travel over the rough surface?

7

A spring is part of a device to propel an object. The constant of the spring is 3.2×10^2 N/m and the mass of the projectile is 8 grams. Before release, the spring is compressed by 6cm. and the projectile is placed in contact with the spring; the spring is then released. What will be the speed of the projectile when the spring reachs equilibrium length?

8

A block resting in the back of a truck begins to slide towards the back as the truck begins moving. The driver is at rest with respect to the truck and concludes that a force is causing the block to move to the rear of the truck. Is he correct? What forces are acting on the block?

9

One end of a unstretched, horizontal spring of stiffness 10N/m is fixed while the other is attached to a block on a frictionless surface. A constant force of magnitude 10N is exerted on the block. What is the speed of the block when it has moved 0.5m?

9 (original)

One end of a unstretched, horizontal spring is fixed while the other is attached to a block on a frictionless surface. A constant force of magnitude 10N is exerted on the block. What is the speed of the block when it has moved 0.5m?

10

A record turntable is coasting (with the motor dis- engaged) at $33^1/_3$ rev/min. when a stack of 10 records suddenly drops on it. Using moments of inertia give an expression for the kinetic energy before and after they drop?

11

A man of mass 80kg lowers himself to the ground from a height of 5 metres by holding onto a rope passed over a pulley and attached to a block of mass 70kg, which is initially at rest. Neglect the mass and friction of the rope and pulley. What is the tension in the rope ?

12

The collision between an automobile and a wall lasts 0.12 sec. The mass of the automobile is 1700kg and the initial and final speeds are v=13.6m/s and v'=-1.3m/s, respectively. What is the time-average force?

13

An enormous turntable rotates about a fixed vertical axis, making one revolutions in 10s. The moment of inertia of the turntable about this axis is $1200kgm^2$. A man of mass 70kg, initially standing at the center of the turntable, runs out along a radius. What is the angular velocity of the turntable when the man is 2m from the center?

14

Two masses m_1= 1.5 kg and m_2= 3.0 kg are connected by a thin string running over a massless pulley. One of the masses hangs from a string; the other mass slides on a 35^o ramp with a coefficent of kinetic friction = 0.40 . What is the acceleration of the masses?

15

A simple manual winch consists of a drum of radius 4cm to which is attached a handle of 25cm. When you turn the handle the rope winds up on the drum and pulls the load. Suppose that the load carried by the rope is 2500N. What force must you exert on the handle to hold the load?

16

A 5 kg block is released from a compressed spring which has a force constant is 120 N/m. After leaving the spring, it travels over a horizontal surface, with a coefficent of friction 0.20, for a distance of 8 metres before stopping. How far was the spring compressed before being released?

17

A rifle bullet of mass 0.01kg strikes and embeds itself in a block of mass 0.99kg which rests on a horizontal frictionless surface and is attached to a coil spring. The impact compresses the spring 10cm. Calibration of the spring shows that a force of 10N is required to compress the spring 1cm. With what velocity did the bullet strike the block?

17 (original)

A rifle bullet of mass 0.01kg strikes and embeds itself in a block of mass 0.99kg which rests on a horizontal frictionless surface and is attached to a coil spring. The impact compresses the spring 10cm. Calibration of the spring shows that a force of 10N is required to compress the spring 1cm. Find the potential energy of the spring?

18

A spring with a force constant k=150 N/m has a released length of 0.15m. What force must be exerted to stretch the spring to twice its length?

19

An empty elevator is an unknown height above the ground and consists of an elevator cage of mass 1200kg connected by a cable, running over a pair of pulleys, to a counterweight 1300kg. Neglect the mass and friction of the cable and pulleys. When the elevator has decended 10 metres how much kinetic energy has the counterweight gained?

20a

A man of mass 80kg lowers himself to the ground from a height 5 metres by holding onto a rope passed over a pulley and attached to a block of mass 70kg, which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. With what speed does the man hit on the ground?

21

A thin rod of mass M and length x hangs from a pivot at its upper end. A ball of clay of mass m and of horizontal velocity v strikes the lower end at right angles and remains stuck (a totally inelastic collision). How high will the rod swing after this collision?

22

A 5kg mass has an initial velocity of 4m/s when it strikes a stationary 3kg mass, and a final velocity of 2m/s. Find the final speed of the 3kg mass and the direction of the masses with respect to the initial direction of the 5kg block?

23

A man of mass 80kg lowers himself to the ground from a height 5 metres by holding onto a rope passed over a pulley and attached to a block of mass 70kg, which is initially at rest. The mass of the man is greater than the mass of the block. Neglect the mass and friction of the rope and pulley. With what speed does the man hit on the ground? (Derive a formula for the speed without using kinematic relations e.g. v2 = u2 + 2as, and evaluate it).

Appendix C

The sorting scores of experiment 1b (Chapter 2)

Sorting by *principle*

Subject number	P
1	0.33
2	0.33
3	0.33
4	0
5	0.25
Average	0.25

Sorting by solution method

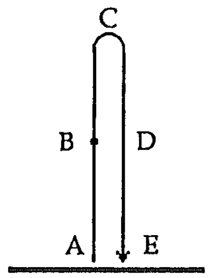
Subject number	P
6	0.20
7	0.38
8	0
9	0
Average	0.145

Sorting by solution

Subject number	P
10	0
11	0
12	0
13	0.17
Average	0.04

Appendix D

The Coin Problem (Clement 1982)



A coin is tossed from point A straight up into the air and caught at point E. On the dot on your paper draw one or more arrows showing the direction of each force acting on the coin when it is at point B. (Draw longer arrows for longer forces).

Appendix E

The sorting scores of experiment 3b (Chapter 4)

Honours class

Sorting by *principle*

Subject number	P
1	0.75
2	0.33
3	0.71
Average	0.60

Sorting by solution

Subject number	P
4	0.25
5	0.71
Average	0.48

Standard class

Sorting by *principle*

Subject number	P
6	0
7	0.25
8	0.57
Average	0.27

Sorting by solution

Subject number	P
9	0.0
10	0.25
11	0.0
Average	0.08

References

Anderson, J.R. (1980) *Cognitive psychology and its implications.* W.H.Freeman & Co.: San Fransisco.

Anderson, J.R. (1987a) Methodologies for studying human knowledge. *Behavioral and Brain Sciences,* 10, 467-505.

Anderson, J.R. (1987b) Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review,* 94, 192-210.

Brown, J. S. & VanLehn, K. (1980) Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science,* 4, 379-426.

Bruner J. S. & Potter M. C. (1964) Interference in visual recognition. *Science,* 144, 424-425.

Chartoff, B. T. (1977) An exploratory investigation utilizing a multidimensional scaling procedure to discover classication criteria for algebra word problems used by students in grades 7-13 (Doctoral dissertation, Northwestern University, 1976). *Dissertation Abstracts International,* 1977, 37, 7006A.

Chi, M.T.H., Feltovich,P.J. & Glaser,R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5, 121-152.

Chi, M.T.H., Glaser,R. & Rees, E. (1981) Expertise in problem solving. In Sternberg, R. (Ed.), *Advances in the psychology of human intelligence.* Lawrence Erlbaum Associates: Hillsdale, N.J.

Chase,W. G. & Simon, H. A. (1973) Perception in chess. *Cognitive Psychology,* 4, 55-81.

Clement, J. (1982) Students' perconceptions in introductory mechanics, *American Journal of Physics,* 50, 66-71.

diSessa, A. A. (1982) Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science,* 6, 37-75.

Ericcson, K. A. & Simon, H. A. (1984) *Protocol Analysis.* M.I.T. Press: Cambridge, Massachusatts

de Groot, A. (1966) Perception and memory versus thought: Some old ideas and recent findings. In Kleinmunz (Ed.), *Problem solving.* Wiley: New York.

Heller, J. I. & Greeno, J. G. (1978) Information processing analyses of mathematical problem solving. In Tyler, R. W. & White, S. H. (chairmen) *Testing, teaching, and learning: report of a conference on research on testing.* U.S. Department of Health, Education, and Welfare, National Institute of Education, Washington, D. C.

Holyoak, K.J. & Koh, K. (1987) Surface and structural similarity in analogical transfer. *Memory and Cognition,* 15, 332-340.

de Jong, T. & Fergusson-Hessler, M. G. M. Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology,* 78, 279-288

Kane, J.W. & Sternheim, M.M.(1983) *Physics* 2nd. Edition.Wiley : New York.

Labudde, P. Reif, F. & Quinn, L. (in press) Facilitation of scientific concept learning by interpretation pocedures and diagnosis. *European Journal of Science Education.* (Copy acquired from authors).

Larkin, J. (1981) Enriching Formal Knowledge. In Anderson, J. R. (Ed.) *Cognitive Skills and Their Acquisition.* Lawrence Erlbaum Associates: Hillsdale,NJ.

Larkin, J., McDermott, J., Simon, D. P. & Simon, H.A. (1980) Expert and novice performance in solving problems. *Science.* 208, 1335-1342.

Leach, C. (1979) *Introduction to Statistics: A nonparametric approach for the social sciences.* Wiley: Chichester, U.K.

Levine, M. (1971) Hypothesis theory and non-learning despite S-R reinforcement contingences. *Psychological Review,* 78, 130-140.

McCloskey, M., Caramazza, A. & Green, B. (1980) Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science,* 210, 1139-1141.

McKendree, J. & Anderson, J. R. (in press) Frequency and practise effects on the composition of knowldge in LISP evaluation. In Carroll, J. M. (Ed.), *Cognitive aspects of human-computer interaction.*

Means, M. L. & Voss, J. F. (1985) Star Wars: A developmental study of expert and novice knowledge structures. *Journal of Memory and Language.* 24, 746-757

Newell, A & Simon, H.A. (1972) *Human problem solving.* PrenticeHall: NJ.

Ohanian, H.C. (1985) *Physics* W.W. Norton & Co.:New York.

Ohlsson, S. (1987) Truth versus appropriateness: Relating declarative to procedural knowledge. In Klahr, D., Langley, P. & Neches, R. (Eds.) *Production System Models of Learning and Development.* M.I.T. Press: Cambridge, Mass.

Posner, M. I.(1973) *Cognition: An introduction.* Scott, Foesman: Glenview,Ill.

Reif, F. (1985) Acquiring an effective understanding of scientific concepts. In West, L.H.T. (Ed.) *Cognitive Structure and Conceptual Change.* Academic Press: Orlando,FL.

Reif, F. (1987) Interpretation of scientific or mathematical concepts: Cognitive issues and instructional implications. *Cognitive Science,* 11, 395-416.3

Reif, F. & Heller, J. I. (1982) Knowledge structure and problem solving in physics. *Educational psychologist,* 17, 102-127.

Sears, F.W., Zemansky,M.W. & Young, H.D. (1982) *University Physics* 6th Edition. AddisonWesley: Reading, Massachusatts.

Scheonfeld, A. H. & Herrmann, D. J. (1982) Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and*

Cognition, 8, 484-494.

Silver, E. A. (1979) Student perceptions of relatedness among mathematical verbal problems. *Journal for Research in Mathematics Education*, 10, 195-210.

Singley, M. K. & Anderson, J. R. (1986) A key-stroke analysis of learning and transfer in text-editing. Manuscript submitted for publication.

Sweller, J. & Gee, W. (1978) Einstellung, the sequence effect, and hypothesis theory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 513-526.